Journal
of
The Franklin Institute
Devoted to Science and the Mechanic Arts

Vol. 270 AUGUST, 1960 No. 2

NEW GYROS FOR OUR SUBMARINES *

BY

FREDERICK D. BRADDON [1]

ABSTRACT

This report describes the present day use of gyro compasses on naval vessels. The nuclear powered submarine with its capability for long sustained submergence and long range cruising, presents new and challenging problems of navigation. Two new types of gyro compass have been introduced by the Navy in recent years to meet the increasing navigational and ordnance requirements of the fleet. Several hundred installations of these equipments have been made over the past four years in both submarines and surface ships. This background of experience serves to establish thoroughly the operating characteristics of these equipments. This talk describes the design features, operating principles, and performance capabilities of the Mark 19 Gyro Compass and its companion instrument, the Mark 23 Gyro Compass. Also described are some of the specialized problems of polar navigation and the techniques employed by the *U.S.S. Nautilus*, the *Skate*, and the *Sargo* in the use of their gyro compasses for their successful polar transits. The Mark 19 Gyro Compass represents the first fully stabilized directional reference with automatic north seeking control, and is the direct forerunner of the Ship Inertial Navigation Systems. The Mark 23 Gyro Compass is unstabilized and indicates only azimuth. It serves for smaller vessels and as an auxiliary compass in combination with the Mark 19. I will try to show the continuing function of the gyro compass in relation to this more complex type of GYRO NAVIGATOR™ now being employed on the new Polaris submarines.

On every naval vessel one will find an instrument known affectionately as the Gyro. This is the compass on which the ship depends to guide it on its way. Each man aboard knows that his safety and his mission depend upon this Gyro every minute of the day. It functions unattended and incessantly for months on end. It is the central datum to which the ship is steered, and to which all navigation, surveillance and ordnance functions are referred.

* Lecture delivered May 4, 1960, at a Joint Meeting of the Philadelphia Science Teachers Association and The Franklin Institute.

[1] Chief Engineer, Marine Division, Sperry Gyroscope Company, Syosset, N. Y.

(Note—The Franklin Institute is not responsible for the statements and opinions advanced by contributors in the JOURNAL.)

The Gyro is not a conspicuous instrument, for it is small and quiet and is generally installed in a protected area deep within the vessel. It requires only electrical connections for its operation. Its output information is transmitted electrically to the remote navigation, surveillance and ordnance stations which need to know true-north and the zenith. This equipment is technically designated as the ship's Master Gyro Compass.

Many forms of gyro compass have been developed and employed by the Navy since the first installation on the *U.S.S. Delaware* in 1911. Changing requirements and changing technology have resulted in a succession of designs that have progressively led to improved accuracies, size reduction and a high degree of operating reliability. It is the purpose of this talk to describe the two currently manufactured Navy type compasses known as the Mark 19 and Mark 23 which are in wide operational use. Figure 1 shows the design and relative size of these equipments.

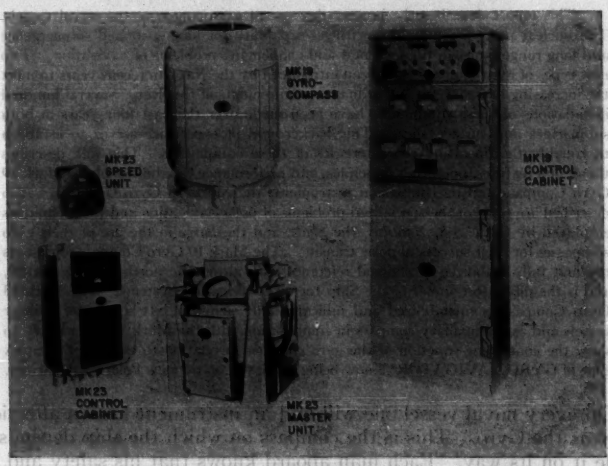

FIG. 1. Composite photos of MK 19 and MK 23 gyro compasses.

The origin of the Mark 19 dates back to World War II when we were confronted with the problem of precisely stabilizing and directing a high resolution shipborne radar designed for aircraft surveillance and air traffic control on carriers. We were amazed to discover the inadequacy of existing shipboard gyro compasses and stable elements for this purpose. Gyro compass accuracy was reasonably adequate for navigation, but transient perturbations in the output data hindered its use

for azimuth stabilization. The stable elements lacked vertical accuracy due to instrumental errors caused by gyro friction, mass shift and inadequate compensation of maneuvering and seaway accelerations. A comprehensive study was initiated to analyze this problem from which evolved the concept of a unitary three-axis gyroscopic element to replace the two separate instruments for azimuth and vertical indication.

Information was collected on the various forms of gyro compass and gyro vertical which were in general use in 1944, or which had been previously developed to the point of successful operation. We catalogued the detailed design and operating characteristics of these equipments under the headings listed in Fig. 2. From these data we

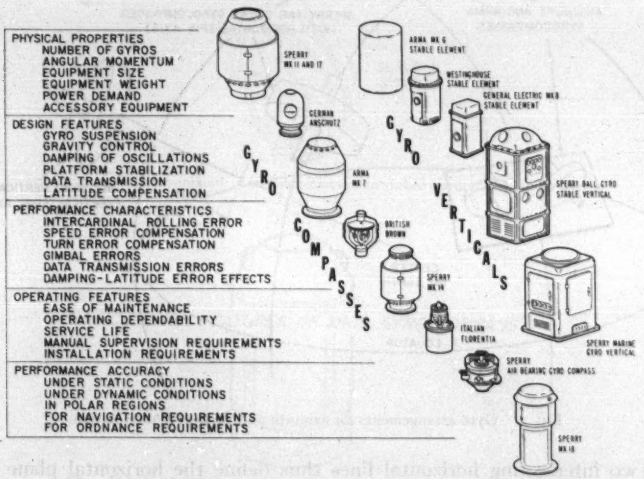

FIG. 2. Evaluation of gyro compasses and stable vertical equipment as of November, 1944.

attempted to analyze and evaluate the theoretical and practical aspects of each design feature. On this basis we established the initial specifications for what was to become the MK 19 Gyro Compass. This tabulation further indicated that all the gyro compasses used gyros with a *horizontal* spin axis. Figure 3 shows that Sperry and Brown used a single gyro in the meridian plane, while Anschutz and Arma used two horizontal gyros linked to subtend the meridian plane. All the gyro verticals (sometimes referred to as stable elements) used a single gyro with a *vertical* spin axis. At this point a very significant observation

was made. *The most precise level reference on a ship is the level attitude of the spin axis of the gyro compass.* Because of the extreme freedom of suspension of the gyro compass about the vertical axis, the gyro is free of perturbations in level. The gyro compass was therefore capable of more accurate indication of tilt in the meridian plane than any of the gyro verticals. From this observation it was deduced that an identical gyro with horizontal spin axis slaved 90° from the meridian gyro, as shown in Fig. 4, would sense tilt in the E-W plane with equal precision.

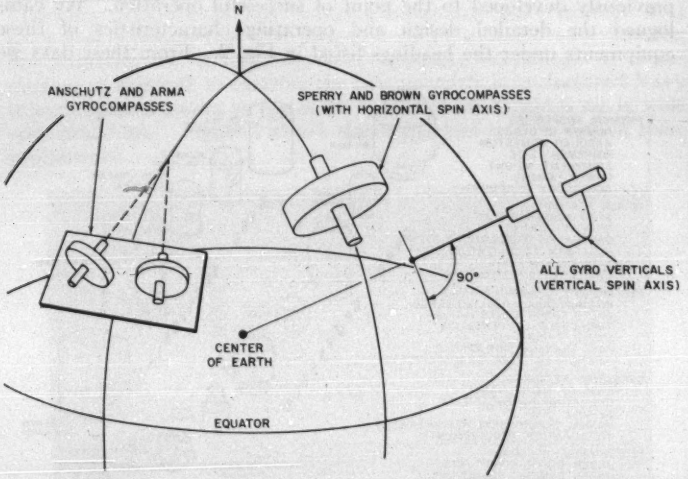

FIG. 3. Gyro arrangements for azimuth and vertical determination.

Two intersecting horizontal lines thus define the horizontal plane to establish a zenith reference and the north seeking control of the meridian gyro continued to point its spin axis to true north. The MK 19 concept thus evolved as a precise and compact single instrument to replace the separate indicators. We originally called this instrument the "Zenith Meridian Indicator," and later conformed to Navy terminology using the designation "Gyro Compass Equipment Stabilized MK 19." This title has tended to de-emphasize the vertical indicating capability of the instrument.

As the result of this study and extensive experimentation and prototype manufacture, the present MK 19 Mod 3 design was established having the basic features listed in Fig. 5. The idea of the composite vertical and azimuth indicator, the use of electrical methods of applying

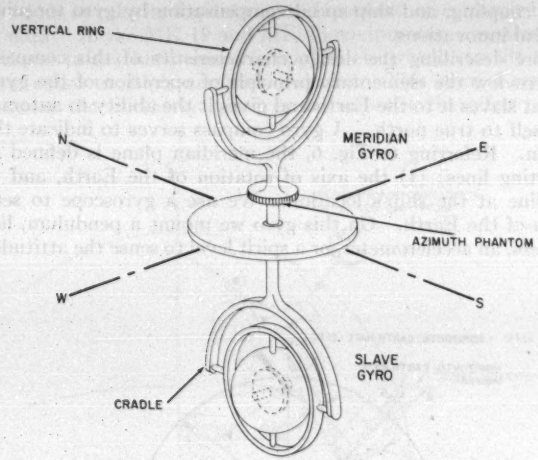

FIG. 4. Concept of unitary three-axis stable element.

BASIC DESIGN FEATURES OF MK19 GYRO-COMPASS

1. TWO 2° OF FREEDOM GYROS, HORIZONTAL SPIN AXIS
2. GYRO SUPPORT BY FLUID EMERSION.
3. NEUTRAL BOUYANCY AND BALANCE OF GYRO UNITS.
4. GYRO CONTROL BY ELECTRIC TORQUERS
5. LEVEL SENSING BY ACCELEROMETERS.
6. SHIP SPEED COMPENSATION BY GYRO TORQUING.
7. DESIGN SYMMETRY IN ALL MECHANICAL STRUCTURE.
8. PORTABLE INTACT THROUGH THE STANDARD SUBMARINE HATCH.
9. SELF CONTAINED SHOCK MOUNTING

FIG. 5.

gravity coupling, and ship speed compensation by gyro torquing were all radical innovations.

Before describing the design characteristics of this compass, it is well to review the elementary principle of operation of the gyro compass that slaves it to the Earth and gives it the ability to automatically align itself to true north. A gyro compass serves to indicate the local meridian. Referring to Fig. 6, the meridian plane is defined by two intersecting lines: (1) the axis of rotation of the Earth, and (2) the plumb-line at the ship's location. We use a gyroscope to sense the rotation of the Earth. On this gyro we mount a pendulum, liquid in containers, an accelerometer, or a spirit level to sense the attitude of the

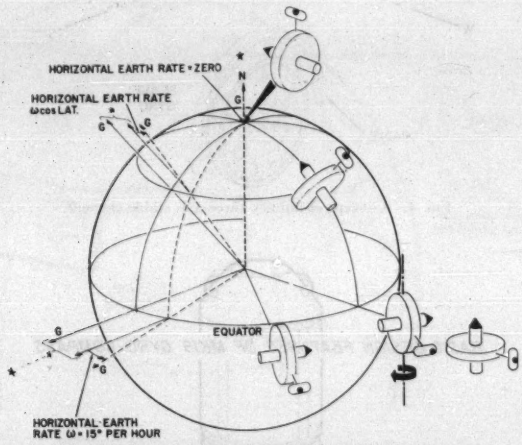

FIG. 6. North-seeking principle of the gyro compass.

gyro to the gravity vector. As the Earth turns under the gyro, which tends to remain fixed in reference to the stars, the spirit level detects a tilt. In the MK 19 and MK 23, an electrical signal is derived from this level and an electrical torque applied to direct the gyro toward the meridian. The gyro thus nulls on the horizontal component of Earth rotation to settle in the meridian plane. A natural error in azimuth indication results from the northerly velocity of the ship. The instrument can sense only the total angular velocity of the Earth and ship. Figure 7 shows this effect to cause the gyro to settle to the west of the meridian for northerly transits. In all previous types of gyro compass, the gyro spin axis was deflected from the meridian by a relatively large amount as a function of the ship's speed, heading, and latitude. This error was computed and corrected in the output data. As a result,

perturbations in velocity log data caused instantaneous transients in the output data. In the MK 19 and MK 23 equipments, a vertical torque proportional to northerly speed is applied electrically to tilt the gyro at a rate equal to ship's northerly velocity. This control completely compensates the northerly speed error, smooths the effect of rough

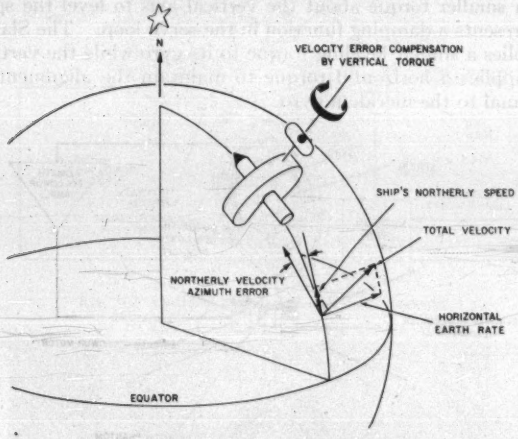

FIG. 7. Northerly speed azimuth error.

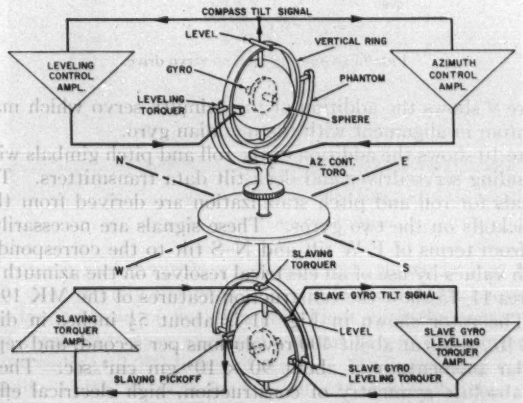

FIG. 8. Gyro leveling and azimuth control.

velocity data, and maintains the gyroscopic element fixed in the meridian plane.

Without too much detail, Fig. 8 shows schematically the composition of the MK 19 Stable Element. The level signal of the meridian gyro is amplified and applies a horizontal torque to direct the system to the meridian as previously described. Simultaneously, the same signal applies a smaller torque about the vertical axis to level the spin axis. This represents a damping function in the servo loop. The Slave Gyro level applies a similar leveling torque to its gyro while the vertical axis pickoff applies a horizontal torque to maintain the alignment of this gyro normal to the meridian gyro.

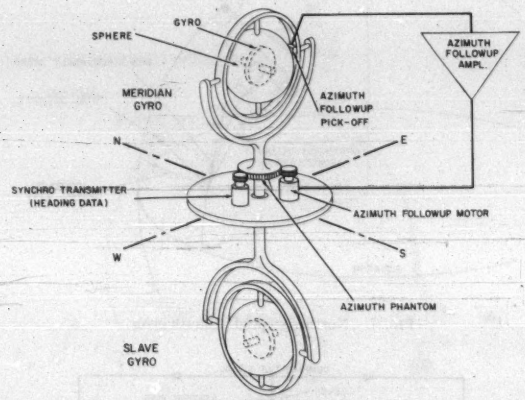

FIG. 9. Azimuth phantom servo drive.

Figure 9 shows the addition of the azimuth servo which maintains the phantom in alignment with the meridian gyro.

Figure 10 shows the addition of the roll and pitch gimbals with their corresponding servo drives and deck tilt data transmitters. The control signals for roll and pitch stabilization are derived from the horizontal pickoffs on the two gyros. These signals are necessarily transformed from terms of E-W tilt and N–S tilt to the corresponding roll and pitch values by use of an electrical resolver on the azimuth drive.

Figures 11–15 show the construction features of the MK 19 Equipment. The rotor shown in Fig. 11 is about $5\frac{1}{2}$ inches in diameter, weighs 8 lb., spins at about 400 revolutions per second, and represents an angular momentum of about 90×10^6 gm cm^2/sec. The design stresses absolute symmetry of construction, high electrical efficiency, and an efficient lubrication system.

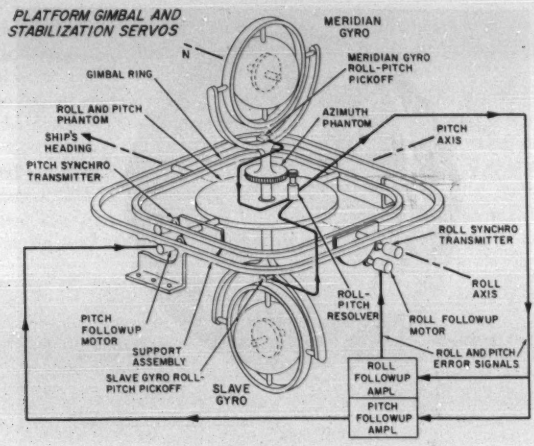

FIG. 10. Platform gimbal and stabilization servos.

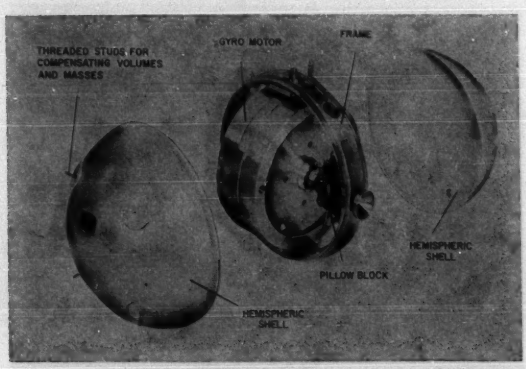

FIG. 11. Gyro motor construction.

The rotor is sealed within a spherical case, shown in Fig. 12. This assembly is accurately balanced under fluid for neutral buoyancy and to bring the centroid of flotation coincident with the center of mass. Many of these units have operated without benefit of maintenance for periods exceeding 12,000 hours of service. An interesting comparison is to consider 12,000 hours of driving your car at 30 mph., thus travelling 360,000 miles, without an oil change or adjustment of any kind! This

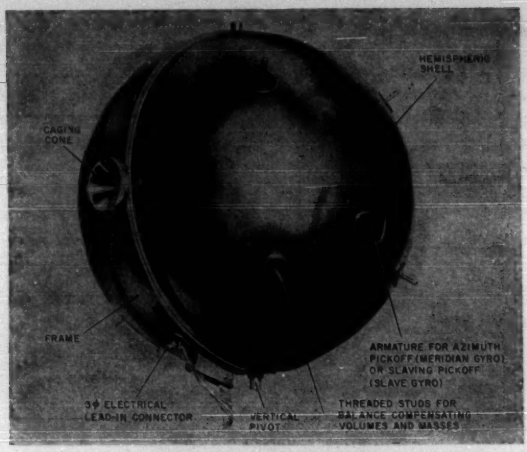
HEMISPHERIC SHELL
CAGING CONE
FRAME
ARMATURE FOR AZIMUTH PICKOFF (MERIDIAN GYRO) OR SLAVING PICKOFF (SLAVE GYRO)
3φ ELECTRICAL LEAD-IN CONNECTOR
VERTICAL PIVOT
THREADED STUDS FOR BALANCE COMPENSATING VOLUMES AND MASSES

FIG. 12. Gyro float assembly.

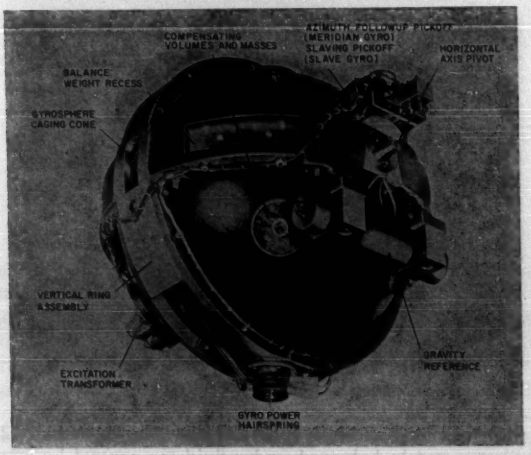

FIG. 13. Gyro gimbal and level assembly.

12,000 hours (or almost 1½ years) of continuous operation at 400 revolutions per second represents 17 billion revolutions of the rotor with a peripheral travel of the rim of 3½ million miles. One of the early installations on an ice breaker logged over 19,000 hours of satisfactory

operation. Two design features contributing to this reliability of operation provide a thermally controlled lubrication system and positive control of the rotor bearing preload.

The sealed gyro unit is then installed in a sensitive gimbal shown in Fig. 13. The electrolytic level, angular pickoff and slip rings are then added to this member. This assembly then mounts on the base of the oil filled enclosure as pictured in Fig. 14. The Slave Gyro is similar in all construction details.

These completed units are fully adjusted and tested separate from the system, and are stored and installed as interchangeable assemblies.

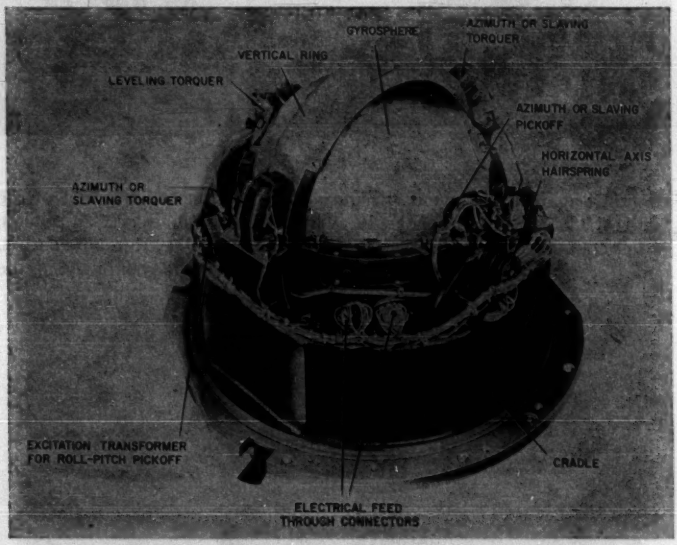

FIG. 14. Meridian gyro mounting.

Figure 15 shows the meridian and slave units installed on the stabilized platform ready for system tests. (Note that the shock mounts are incorporated within the binnacle and a mechanical pantograph linkage serves to provide positive alignment of the stabilization axes.)

The very slow response characteristics inherent in the gyro control system make it mandatory to avoid any interruption in operation of the equipment. Even a momentary interruption will destroy system accuracies for hours to follow. Provision is made in equipment design

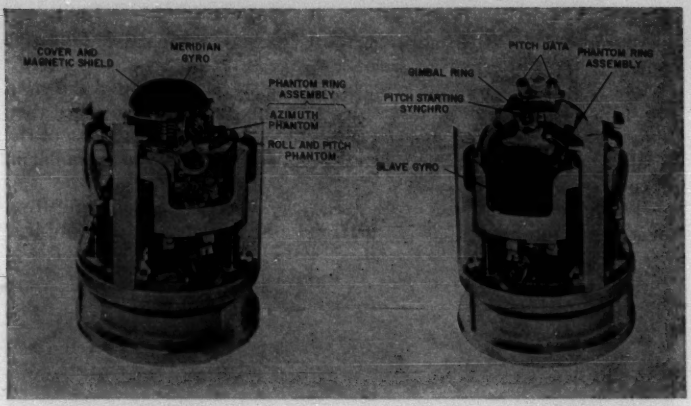

FIG. 15.

to guard against this event. Figure 16 shows the construction of the Control Cabinet of the MK 19, which houses the servo amplifiers, power converters, computers, electro-mechanical integrators, indicators, and test circuits. Magnetic amplifiers are used for reliability in the

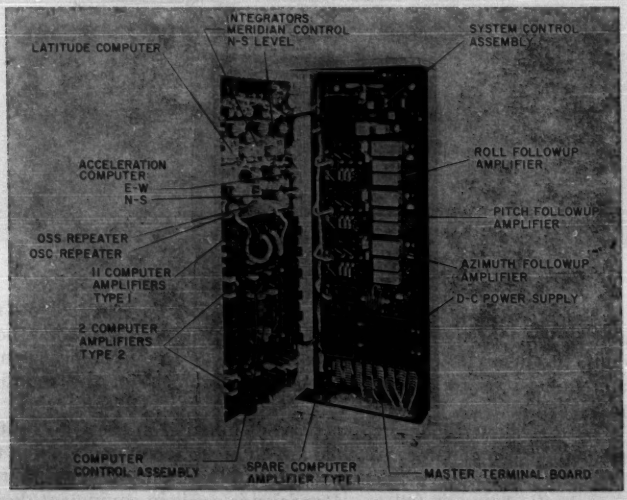

FIG. 16. MK 19 control cabinet.

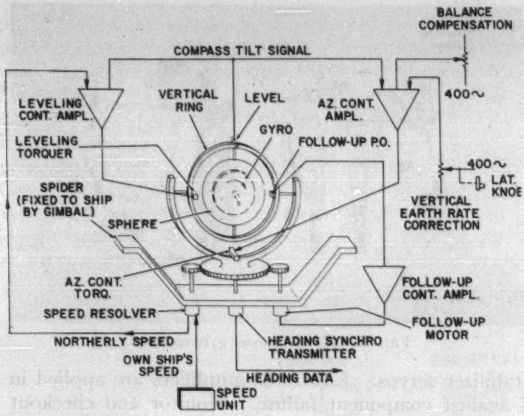

FIG. 17. MK 23 gyro compass gyro and gimbal arrangement.

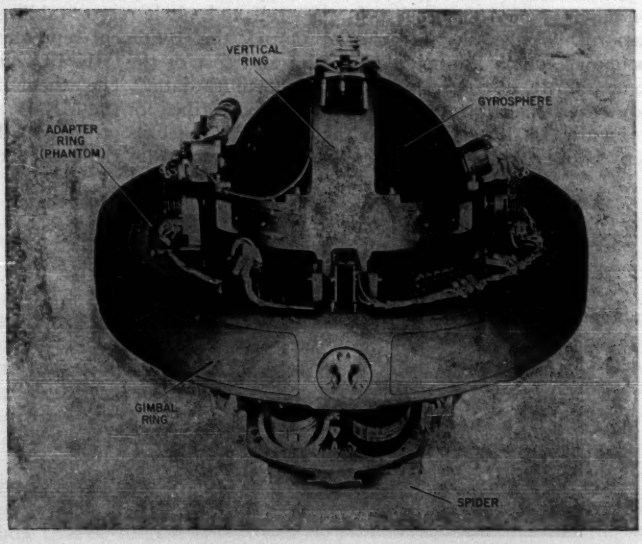

FIG. 18. MK 23 pendulous platform.

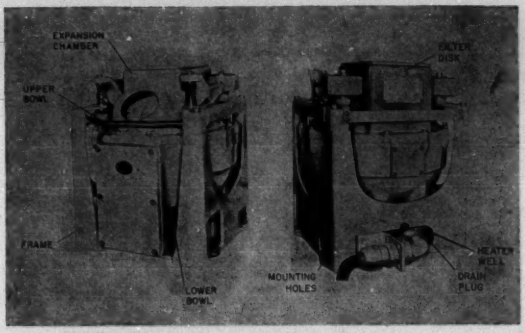

FIG. 19. MK 23 master gyro compass.

critical stabilizer servos. Electronic amplifiers are applied in parallel to guard against component failure. Monitor and checkout circuits permit failure or replacement of any component without interruption to system operation.

The MK 23 Gyro Compass, represented in Fig. 17, incorporates all of the features of the MK 19 meridian gyro, but substitutes a simple pendulous gimbal suspension for the roll and pitch servo stabilized platform. With other design compromises, the equipment represents about one fourth the complexity, size and cost of the MK 19 Gyro Compass. It, of course, provides only azimuth data. The schematic shows the same control features previously described for the MK 19.

Figures 18, 19, and 20 show the construction details of the MK 23 Gyro Compass. Its operating reliability is favored by the avoidance

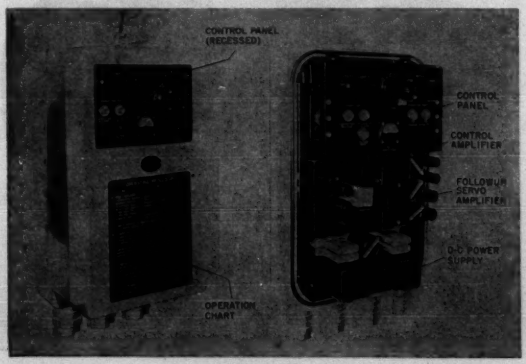

FIG. 20. MK 23 control cabinet.

of the critical stabilizing servos, by its over-all simplicity, and by its low demand on ship's electrical power. When used as an emergency back-up for MK 19 installations, care should be exercised to see that the MK 23 operates from an independent source of ship's power. Interruptions in the switching of ship's power is the most frequent cause of error in compass operations.

The generally accepted performance accuracies of the MK 19 and MK 23 Gyro Compasses exceed the design specifications by an appreciable amount. Figure 21 represents a summary tabulation showing the static and dynamic performance capabilities of these equipments. To generalize, we might classify the MK 19 as accurate in azimuth to

	AZIMUTH		AZIMUTH		ROLL AND PITCH	
FACTORY TEST LIMITING VALUES (DEGREES)	SPEC	ACTUAL	SPEC	ACTUAL	SPEC	ACTUAL
SETTLING POINT REPEATABILITY	0.2	0.2	0.15	0.10	0.07	0.03
TURNTABLE VARIOUS HEADINGS	0.7	0.3	0.15	0.10	0.07	0.03
LOW ANGLE SCORSBY	1.0	0.5	0.20	–	0.15	0.04
HIGH ANGLE SCORSBY	2.0	0.7	0.40	0.20	0.20	0.05
SHIPBOARD TEST RMS VALUES (DEGREES)						
DOCKSIDE REPEATABILITY	–	0.2	–	0.05	–	0.02
LIGHT SEA NO MANEUVERS	–	0.2	–	0.05	–	0.03
MODERATE SEA NORMAL MANEUVERS	–	0.5	–	0.10	–	0.03
HEAVY SEA RADICAL MANEUVERS	–	1.0	–	0.20	–	0.05

FIG. 21. MK 19 and MK 23 gyro compass performance accuracy (at 40° latitude).

one tenth degree when operating in mid latitudes under normal shipboard conditions. The corresponding accuracy of the MK 23 is about one half degree. The .05° azimuth repeatability of the MK 19 is a significant figure in establishing the quality of the instrument. This value indicates that the gyro is consistently, day after day, measuring Earth rates of 11.7°/hour × sin .05°, or .01°/hour.

To the gyro engineer and navigator, the performance of the MK 19 and 23 Gyro Compasses on the *Nautilus* polar transit was almost as spectacular as the mission itself! The actual compass log entries show that on August 1, 1958 at 5:00 P.M. and latitude 71–51.8, the submerged ship was pointed to the North Pole. With a MK 19 heading of

000 and in automatic steering, the ship cruised for seven hours that day, all the next day, and for 23½ hours the following day, to latitude 88.46 without alteration of course. With the Pole indicated as less than 90 miles ahead, a small course correction of 3° was made to correct an indicated offset of only 4½ miles after 1100 miles of transit! It is equally impressive to note that the divergence of the MK 23 from the MK 19 never exceeded ±1°. This circumstance was especially reassuring to the operation.

The pictorial representation of this voyage, shown in Fig. 22, (and with acknowledgment to *National Geographic*) shows that after a total submerged transit of over 96 hours and 1,800 miles, the initial celestial position fix checked within 10 miles of the dead reckoning position.

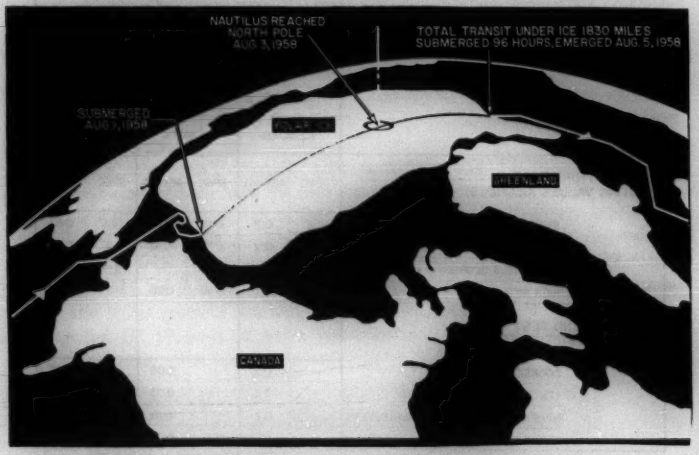

FIG. 22. *Nautilus* polar transit.

The *Nautilus* expedition was planned in a manner to favor gyro compass operation with straight courses, no maneuvering accelerations, and minimum time of transit. MK 19 performance on the *Skate* and *Sargo* was equally satisfactory below 80° latitude, but showed fairly high drift rates while maneuvering for days in the extreme polar region.

Although the MK 19 Gyro Compass was not designed to be operated as a Directional Gyro for grid-type navigation in the polar region, the equipments on the *Nautilus*, *Skate* and *Sargo* have been modified for such use. The present design is not ideally suited to this type operation because of the mass shift effects of the bubble in the electrolytic level and the variable sensitivity of this device. A Navy sponsored development program now in progress will greatly increase the Directional

Gyro stability of the MK 19 for polar navigation. Figure 23 illustrates the bubble level presently used and the force feedback pendulum which will replace the spirit level. In the spirit level, the variable wetting of the electrode, with tilt, alters the balance of the resistance bridge circuit. A clean signal with high tilt sensitivity is provided by a small and simple, reliable device. Thermal sensitivity, shock sensitivity, mass shift effects, and response lag are, however, considered to limit system performance accuracies, particularly in the Directional Gyro mode of operation.

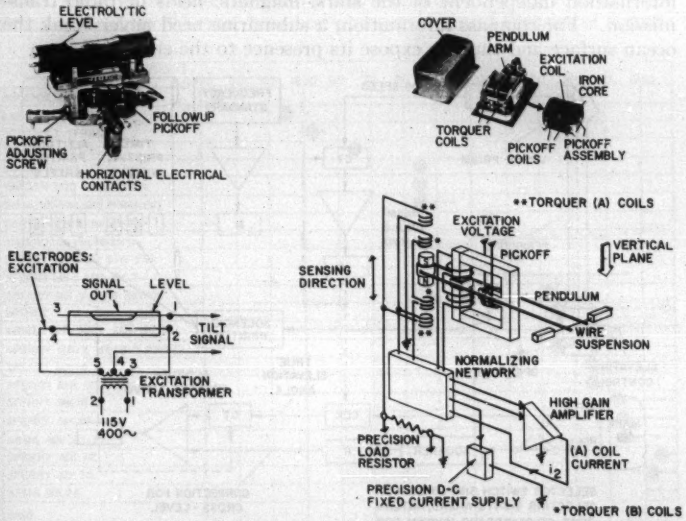

FIG. 23. Liquid level *versus* torque feedback accelerometer.

The force feedback pendulum is somewhat more complex, but is expected to avoid all of the operating limitations of the spirit level and improves both the static and dynamic accuracies of the MK 19 System. This modification is equally applicable to the MK 23 System. Other features in this redesign of the MK 19 Gyro Compass will further simplify the system and lead to improved performance accuracies.

The requirement for precise zenith and azimuth information, with absolute accuracy and stability for satisfying the increasing demands of ordnance and navigation, represents a continuing need within the Navy for the gyro compass. In this talk I want to help you to fully understand those features that make the gyro compass ideally suited to the normal navigational needs of the nuclear submarine.

The gyro compass is automatic in its alignment to the meridian plane through the Earth's polar axis. It requires only to feel the force of gravity and to sense the daily rotation of the Earth about its polar axis. These two natural quantities persist in all parts of the Earth and for all time. They exist inside the vessel as well as outside, and penetrate the depths of the ocean. They are unaffected by daylight or nightfall, or by the seasons of the year. They cannot be altered or obscured by weather or by man-made interference. The gyro compass thus provides our ships with secure, continuous and precise compass information independent of the stars, magnetic fields or radio transmission. For compass information, a submarine need never break the ocean surface and thereby expose its presence to the enemy.

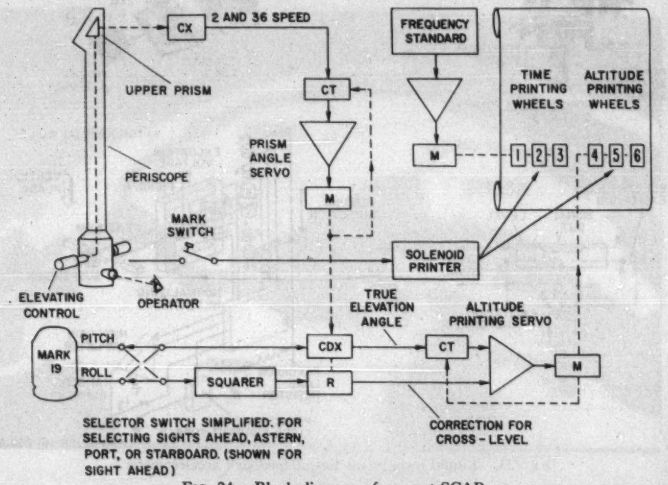

FIG. 24. Block diagram of present SCAR.

Navigational accuracy of dead reckoning is primarily limited by the lack of means to measure speed precisely over the bottom. A ship's heading error of 0.1° represents an offset from course of less than 2 miles in position after 1000 miles of transit. If we could measure groundspeed precisely, this 2 miles would be our position error. Today with uncharted ocean currents and limitations in measuring speed through the water, an uncertainty of 1 knot in estimating bottom speed can develop position errors up to 30 miles or more in a 1000-mile mission. Some day we will find a secure means to measure ground speed.

The MK 19 made available for the first time in a submarine a precise zenith reference. A logical and relatively simple outgrowth of

this program was to bring this information into the periscope so that star altitude measurements could be made for a celestial position fix. This equipment constitutes a periscopic sextant known as the *S*ubmarine *C*elestial *A*ltitude *R*ecorder (SCAR), shown in Fig. 24. This equipment is installed on a number of our advanced submarines and for the first time permits the vessel to establish its position by celestial reference without surfacing.

To place the MK 19 and MK 23 program in perspective I have tabulated a few of the mileposts in the long history of gyro instrumentation. We are told that the unique characteristics of the gyroscope were first analyzed by Newton, based on observations and theories of astron-

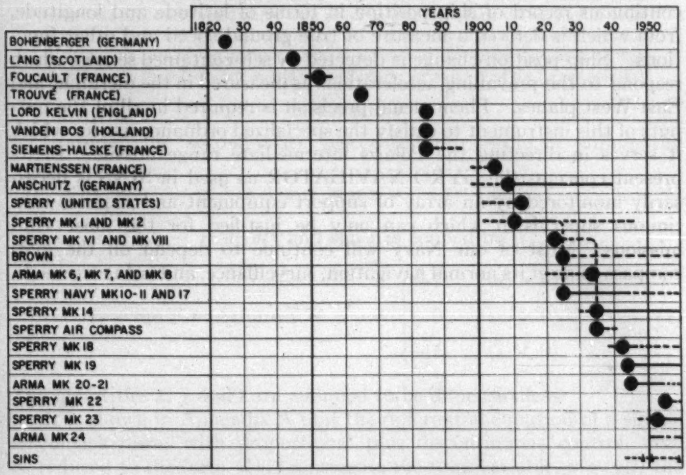

FIG. 25. Chronology of gyro navigation.

omers relating to the cyclic precession of the equinoxes. Figure 25 shows that Bohenberger in 1827 is credited with the conception of the gyroscope as a mechanical device. Lang, about 1846, suggested its possible use as a gyro compass. It remained for Foucault in 1852 to demonstrate this effect and expound the theory of the device. Trouvè, Kelvin, van den Bos and Siemens contributed significant experimentation from 1865 to 1885. Martienssen in 1905 developed, for the first time, an accurate mathematical analysis of the gyro compass. About five years later Anschutz in Germany and Elmer Sperry in New York produced the first practical shipboard equipments.

Following the first Navy gyro compass installation in March, 1913

the Navy has continued to purchase various types of navigational and ordnance gyro compasses. From the early twenties until recent years, Arma was an alternate supplier to the Navy. Figure 25 shows that only MK 14, 19, and 22 types are in current production. The approximate number of Navy ship installations for each type is shown. The total over a span of 50 years represents something over 10,000 equipments.

The very complex *S*hip *I*nertial *N*avigation *S*ystem (SINS), as now being produced for our new fleet of Polaris submarines, represents a new phase in ship navigation. This equipment incorporates the geographic directional reference system of the stabilized gyro compass producing precise zenith and azimuth data. In addition, it provides a continuous record of ship position in terms of latitude and longitude, from which is derived a measure of true ground speed and other functions. Ship position change is detected by self-contained sensors which respond to the prevailing accelerations as measured in the meridian and East-West planes. Phenomenal precision is required in all of the outputs of this instrument to satisfy the specialized ordnance use for which it serves in directing the Polaris intermediate range missile. In its present concept, the GYRO NAVIGATOR as used in SINS is necessarily monitored by an array of support equipment and requires continuous supervision which can only be justified for the most vital missions. Most of our Navy will continue to depend on the gyro compass to meet its normal navigation, surveillance, and ordnance needs.

HARMONIC AND SUBHARMONIC RESPONSE OF AN ON-OFF CONTROL SYSTEM TO SINUSOIDAL INPUTS *

BY

B. A. FLEISHMAN [1]

ABSTRACT

We derive here periodic solutions, of harmonic and subharmonic type, of a differential equation governing the operation of a relay servomechanism which is subjected to a sinusoidal input $y_0 \sin \omega t$. Each solution is a "formal" solution in the sense that it is constructed under an a priori assumption about the manner in which it changes sign, and this assumption can be verified only after one has the solution. A significant part of the analysis, then, is devoted to the determination of suitable conditions on the parameters y_0 and ω which insure the validity of the formal solution.

INTRODUCTION

Consider the second-order piecewise-linear ordinary differential equation

$$\ddot{x} + x + \operatorname{sgn} x = -\ddot{y} \tag{1}$$

where $\cdot = d/dt$, $\cdot\cdot = d^2/dt^2$, and sgn x is defined by

$$\operatorname{sgn} x = \operatorname{signum} x = \frac{x}{|x|} = \begin{cases} +1 & (x > 0) \\ 0 & (x = 0) \\ -1 & (x < 0) \end{cases} \tag{2}$$

The quantities x, y and t are assumed to be dimensionless.

It is shown in Appendix A that the differential equation of a simple servomechanism, with proportional plus discontinuous control, may be reduced to the form (1); t represents the time variable, $y = y(t)$ the (known) input function, and $x = x(t)$ the (unknown) error function, that is, deviation of output from input.

The basic analytical problem associated with any control system is to predict the error $x(t)$ (or, equivalently, the output $z(t) = x(t) + y(t)$) as a function of time, when the input function $y(t)$ and the initial error and error rate are prescribed. Because it is not linear, Eq. 1 cannot be solved exactly for arbitrary inputs $y(t)$. Most analyses of nonlinear control systems deal with step ($y \equiv$ constant) and ramp ($dy/dt \equiv$ con-

* This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract Number AF 49(638)-514. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1] Associate Professor of Mathematics, Rennselaer Polytechnic Institute, Troy, N. Y.

stant) inputs, principally by phase-plane techniques (for example, see (1)[2]). Except by numerical means, very little has been done with other continuously varying inputs $y(t)$, particularly for differential equations with discontinuous nonlinearities, where such classical techniques as iteration and perturbation are inapproprate.

In this paper sinusoidal inputs are considered. Even so, because (1) is not linear, we cannot find an exact solution for arbitrary initial data. Rather, we study only periodic solutions. It is shown that when $y(t)$ is a sine function, the system governed by (1) is capable of steady-state periodic response, of harmonic or subharmonic type, depending upon the initial conditions and the values of the parameters in the input function. A harmonic solution is one with the same period as the input, while a subharmonic response is one whose smallest period is an integral multiple of that of the input. These periodic solutions are obtained by exploiting the piecewise-linear nature of Eq. 1.

The occurrence of subharmonic oscillations is a significant feature of nonlinear systems not shared by linear ones. If its amplitude is large, such an oscillation may have destructive effects. Several writers have considered this problem. West and Douce (2) have analyzed subharmonic responses in nonlinear control systems by means of the describing-function method. Very recently Ogata (3) has studied the phenomenon analytically and experimentally (by use of an analog computer) in a group of systems in which the output of the nonlinear element feeds into linear elements.

The analysis presented here deals with differential equation (4) below; in this system the discontinuous and proportional control elements are in parallel, that is, they act simultaneously on the error signal $x(t)$. We are able to derive explicitly, and verify, the periodic solutions, pointing the way to a study of periodic responses to general periodic (not necessarily sinusoidal) inputs. This extension will be carried out in a future investigation. One of the principal results of this paper is the exact analytical form, in the case of the control system studied here, of the subharmonic oscillations.

Since we only consider inputs of the form

$$y(t) = y_0 \sin \omega t, \tag{3}$$

where the amplitude y_0 and frequency ω are constant, Eq. 1 may be written

$$\ddot{x} + x + \operatorname{sgn} x = y_0 \omega^2 \sin \omega t. \tag{4}$$

Clearly, once we find any particular solution $x(t)$ of Eq. 4, representing the error as a function of time, we shall also have the output, $z = x + y = x + y_0 \sin \omega t$.

[2] The boldface numbers in parentheses refer to the references appended to this paper.

The harmonic case is treated in Part I. First a "formal" steady-state solution of Eq. 4 with period $2\pi/\omega$ is constructed, then restrictions are imposed on the parameters y_0 and ω to insure that this periodic function is indeed a solution of the differential equation. In Part II, by suitable modification of the harmonic solution, we obtain a formal subharmonic solution (of period $2m\pi/\omega$, m an integer), and again we place restrictions on the parameters to insure the validity of this formal solution. The principal results are contained in theorems at the end of Parts I and II, and in Part III, "Concluding Remarks."

I. HARMONIC CASE

1. Derivation of Formal Solution

A harmonic solution of Eq. 4 can be constructed as follows. If two functions $x_f(t)$ and $x_1(t)$ satisfy the respective equations

$$\ddot{x}_f + x_f = y_0\omega^2 \sin \omega t \tag{5}$$

and

$$\ddot{x}_1 + x_1 = -\operatorname{sgn} x_1 \tag{6}$$

then upon adding these equations we see that the function

$$x_h(t) = x_f(t) + x_1(t) \tag{7}$$

is a solution of Eq. 4, provided that $\operatorname{sgn} x_h = \operatorname{sgn} x_1$, that is, provided that x_h and x_1 always have the same sign.

Since there is no damping in the system, let us suppose that x_f, x_1 and x_h are all in phase (or 180° out of phase) with $\sin \omega t$. The particular solution of Eq. 5 with this property is just

$$x_f(t) = \frac{y_0\omega^2}{1 - \omega^2} \sin \omega t. \tag{8}$$

As for x_h and x_1, they must vanish at $t = k\pi/\omega$, $k = 0, 1, 2, \cdots$. Furthermore, since they are to have period $2\pi/\omega$ and be in phase with $\sin \omega t$, it is not unreasonable to assume that they maintain one sign during the half-cycle $0 < t < \pi/\omega$, the opposite sign during the next half-cycle, etc. To be specific, we assume x_1 (and therefore x_h) > 0 during the first half-cycle.

Thus, in view of Eq. 6 we seek a periodic solution $x_1(t)$ of the equations

$$\ddot{x} + x = -1 \qquad (x > 0) \tag{9a}$$

$$\ddot{x} + x = +1 \qquad (x < 0) \tag{9b}$$

having the period $2\pi/\omega$ and satisfying

$$x(0) = 0, \qquad \dot{x}(0) = v_0 > 0, \tag{10}$$

where the constant v_0 remains to be determined. We can gain our objective most elegantly by way of a geometrical discussion.

Multiplying both sides of Eq. 9a by $\dot{x}$ and integrating with respect to t, we get

$$\dot{x}^2 + (x + 1)^2 = \text{constant}. \tag{11a}$$

In other words, all trajectories of Eq. 9a in an x, $\dot{x}$-plane (phase plane) are arcs of circles centered at $x = -1$, $\dot{x} = 0$; if $x(0) = x_0$ and $\dot{x}(0) = v_0$, the solution of (9a) with these initial values is represented by the circular

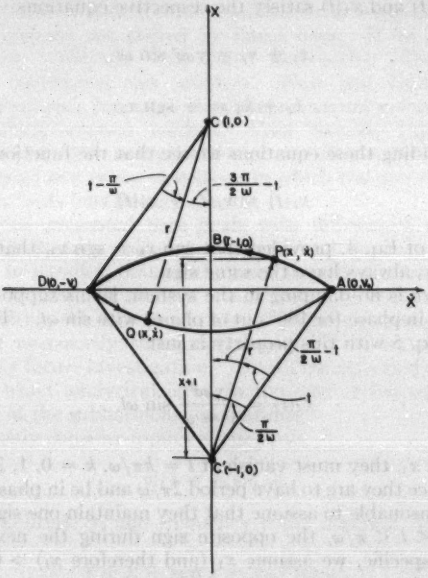

FIG. 1. Construction in phase plane of solution of Eqs. 9 with period $2\pi/\omega$.

arc centered at $(-1, 0)$ which passes through the point (x_0, v_0). When the positive x- and $\dot{x}$-axes extend, respectively, upward and to the right (as in Fig. 1), an arc is traversed in the counterclockwise direction about its center as t increases, for when $\dot{x} = dx/dt$ is positive x must increase

with increasing t. Similarly, all phase-plane trajectories of Eq. 9b are arcs of circles

$$\dot{x}^2 + (x - 1)^2 = \text{constant} \tag{11b}$$

centered at $x = +1$, $\dot{x} = 0$, again always traversed in the counterclockwise direction.

Now note that Eqs. 9 contain restrictions on the sign of x: Eq. 9a applies only when x is positive, Eq. 9b when it is negative. Thus a circular arc of type (11a) may be a portion of a trajectory only while the trajectory is in the upper half-plane. When such an arc, as part of a trajectory, intersects the $\dot{x}$-axis the trajectory continues beyond this point of intersection as an arc of type (11b), until this arc in turn intersects the $\dot{x}$-axis, etc.

A final remark on the nature of all trajectories of Eqs. 9 concerns the time it takes for a given circular arc to be traversed. If Eqs. 11 are written in the form

$$\frac{dx}{dt} = \dot{x} = [\text{const.} - (x \pm 1)^2]^{1/2}$$

and the variables are separated, we see immediately upon integrating that the time elapsed during the traversal of any (circular arc) trajectory is just equal to the angle subtended by the arc at its center.

We are now in a position to construct the required solution $x_1(t)$ of Eqs. 9. First of all, because of the initial conditions (10), the trajectory (see Fig. 1) issues from the point $A(0, v_0)$, following the circle

$$\dot{x}^2 + (x + 1)^2 = r^2 = v_0^2 + 1$$

through the point $B(r - 1, 0)$ on the x-axis to $D(0, -v_0)$. At this point the trajectory switches over to the circle

$$\dot{x}^2 + (x - 1)^2 = r^2,$$

which it follows back to the point $A(0, v_0)$, where it switches again to the original circular arc, and so on. The motion represented by this trajectory is clearly periodic, with quarter-period equal to the angle subtended by the arc AB at its center of curvature $C'(-1, 0)$. Thus, for the period to have the value $2\pi/\omega$, we need merely require that angle $AC'B$ equal $\pi/2\omega$, or that

$$v_0 = \tan \frac{\pi}{2\omega}. \tag{12}$$

To find the form of the function $x_1(t)$, let $P(x, \dot{x})$ be the point on the upper arc representing the state of the system at time t, $0 < t < \pi/\omega$.

Since angle $AC'P = t$, angle $PC'B = \frac{\pi}{2\omega} - t$; therefore, since $|C'P| = r$,

$x + 1 = r\cos\left(\frac{\pi}{2\omega} - t\right)$ and

$$x = x_1(t) = -1 + r\cos\left(\frac{\pi}{2\omega} - t\right) \qquad \left(0 \leq t \leq \frac{\pi}{\omega}\right) \tag{13a}$$

where

$$r = \sqrt{v_0^2 + 1} = \sec\frac{\pi}{2\omega}. \tag{14}$$

Now let $Q(x, \dot{x})$ be a point on the lower arc, corresponding to a value of t in the interval $\pi/\omega < t < 2\pi/\omega$. Since the point $D(0, -v_0)$ corresponds to $t = \pi/\omega$ and angle $DCC' = \pi/2\omega$, angle $DCQ = t - \pi/\omega$ and

$$\text{angle } QCC' = \frac{\pi}{2\omega} - \left(t - \frac{\pi}{\omega}\right) = \frac{3\pi}{2\omega} - t.$$

Therefore the x-coordinate of Q is given by

$$x = x_1(t) = 1 - r\cos\left(\frac{3\pi}{2\omega} - t\right) \qquad \left(\frac{\pi}{\omega} \leq t \leq \frac{2\pi}{\omega}\right). \tag{13b}$$

Equations 13, which represent $x_1(t)$ for $0 \leq t \leq 2\pi/\omega$, must be slightly modified to yield analytic expressions for $x_1(t)$ for all $t \geq 0$. Because of the periodicity, the point $A(0, v_0)$ represents the state of the system at times $t = 0,\ 2\pi/\omega,\ 4\pi/\omega, \cdots$, and $B(0, -v_0)$ at $t = \pi/\omega,\ 3\pi/\omega, 5\pi/\omega, \cdots$; similarly, the upper arc in Fig. 1 is the trajectory during intervals $2k\pi/\omega \leq t \leq (2k+1)\pi/\omega$, $k = 0, 1, 2, 3, \cdots$ and the lower arc during intervals $(2k+1)\pi/\omega \leq t \leq (2k+2)\pi/\omega$. Thus, we need merely increase by $2k\pi/\omega$ the arguments of the cosine in Eqs. 13, to obtain the following representations of $x_1(t)$ valid for all $t \geq 0$:

$$x_1(t) = \begin{cases} -1 + r\cos\left(\dfrac{(4k+1)\pi}{2\omega} - t\right) & \left(\dfrac{2k\pi}{\omega} \leq t \leq \dfrac{(2k+1)\pi}{\omega}\right) \\ 1 - r\cos\left(\dfrac{(4k+3)\pi}{2\omega} - t\right) & \left(\dfrac{(2k+1)\pi}{\omega} \leq t \leq \dfrac{(2k+2)\pi}{\omega}\right) \end{cases} \tag{15}$$

$$k = 0, 1, 2, \cdots$$

2. Restrictions on Parameters

The preceding construction of $x_1(t)$ requires $\omega > 1$, for the angle π/ω subtended by the arc $\widehat{ABD}$ at $C'(-1, 0)$ is always less than π. If $\omega < 1$, solutions $x_h(t)$ may be found by assuming x_1 and x_h 180° out of phase, but we restrict our attention here to harmonic solutions of Eq. 4 with $\omega > 1$, principally because the mode of construction and the form of the results may be conveniently modified for treatment of the subharmonic case.

A further restriction is needed on the parameters y_0 and ω. It has been assumed above that $x_h(t) = x_f(t) + x_1(t) > 0$ for $0 < t < \pi/\omega$, that is,

$$x_f(t) = \frac{y_0\omega^2}{1 - \omega^2} \sin \omega t > -x_1(t) \qquad \left(0 < t < \frac{\pi}{\omega}\right). \tag{16}$$

Since x_1 is positive in this interval, the simplest sufficient condition for Eq. 16 is that y_0 be negative, in which case, since $\omega > 1$, $x_f(t)$ will be positive. Then $x_f(t)$ and $x_1(t)$ are in phase, but the forcing function $y_0\omega^2 \sin \omega t$ is 180° out of phase with them. It is a curious but well-known phenomenon in linear systems that the forced oscillation is respectively in phase, or 180° out of phase, with the applied force according as the natural frequency is greater or less than the frequency of the applied force.

In case y_0 is positive, x_f will be negative during the first half-period, and then appropriate conditions must be imposed on the parameters to insure $x_h = x_f + x_1 > 0$ during this time interval. In Appendix B, for example, it is shown that, for $1 < \omega < 2$, $x_1(t) > -x_f(t)$ when $0 < t < \pi/\omega$, provided y_0 is less than some positive quantity depending on ω. Possible improvements of the result are also mentioned in this Appendix.

The principal results of the preceding analysis are contained in

Theorem 1. Suppose $\omega > 1$. Then if $y_0 < 0$, the function $x_h(t) = x_f(t) + x_1(t)$ is a harmonic solution of Eq. 4 (that is, one with period $2\pi/\omega$), where $x_f(t)$ and $x_1(t)$ are given by Eqs. 8 and 13 (or Eq. 15), respectively. If $y_0 > 0$, $x_h(t) = x_f(t) + x_1(t)$ is a harmonic solution provided the values of $\omega(> 1)$ and y_0 are such that $x_f(t) + x_1(t) > 0$ for $0 < t < \pi/\omega$.

The theorem refers to a periodic solution x_h which is positive during the first half-period. This assumption about the sign of x_h was made only for convenience; it is clear that if x_1 and x_f are multiplied by -1, $x_h = x_1 + x_f$ will again be a harmonic solution, which is, however, *negative* during the first half-period.

II. SUBHARMONIC CASE

1. *Formal Subharmonic Solution*

We are now interested in periodic solutions of Eq. 4, of least period $2m\pi/\omega$, $m = 2, 3, \cdots$. As in the harmonic case, we prove that such solutions exist by constructing them.

Let m be any positive integer greater than 1. Again we seek a solution of the form $x_m(t) + x_f(t)$, where $x_f(t)$ is given by Eq. 8 and $x_m(t)$ is a solution of Eq. 9 with period $2m\pi/\omega$ (obtained in exactly the same way as was $x_1(t)$ in the harmonic case). Since $x_f(t)$ also has the period $2m\pi/\omega$ (it has all periods $2n\pi/\omega$, $n = 1, 2, 3, \cdots$), the function $x_{sm} = x_m + x_f$ will have the required period. (The subscript sm indicates "subharmonic of order m.")

Since for $\omega > 1$, Eqs. 13 furnish a solution of $\ddot{x} + x + \operatorname{sgn} x = 0$ of period $2\pi/\omega$, a solution of period $2m\pi/\omega$ is obtained merely by replacing ω by ω/m in Eq. 13. Thus

$$x_m(t) = \begin{cases} -1 + \sec\dfrac{m\pi}{2\omega}\cos\left(\dfrac{m\pi}{2\omega} - t\right) & \left(0 \le t \le \dfrac{m\pi}{\omega}\right) \\ 1 - \sec\dfrac{m\pi}{2\omega}\cos\left(\dfrac{3m\pi}{2\omega} - t\right) & \left(\dfrac{m\pi}{\omega} \le t \le \dfrac{2m\pi}{\omega}\right) \end{cases} \tag{17}$$

when $\omega/m > 1$, or $\omega > m$. The specification of the function for all $t \geqslant 0$ would be given by equations analogous to Eq. 15. Clearly x_m vanishes at $t = km\pi/\omega$, $k = 0, 1, 2, \cdots$, and it is positive for $0 < t < m\pi/\omega$, negative for $m\pi/\omega < t < 2m\pi/\omega$, and so on. It is worth noting that the initial conditions giving rise to solution (17) are $x_m(0) = 0$, $\dot{x}_m(0) = v_0$, where

$$v_0 = \tan\frac{m\pi}{2\omega}. \tag{18}$$

2. *Restrictions on Parameters*

We now know that $x_{sm} = x_m + x_f$ has the required periodicity property, also that both x_m and x_f vanish at $t = 0, m\pi/\omega, 2m\pi/\omega, \cdots$, but we must make sure, as in the harmonic case, that $x_{sm} = x_m + x_f$ actually satisfies the differential equation (4). In other words, the parameters must be subjected to conditions which will insure that x_{sm} is positive during the first half-period, $0 < t < m\pi/\omega$, negative during the second, etc. The question is more complicated here than in the harmonic case because $x_f(t)$ has $(m - 1)$ zeroes in the interval $0 < t < m\pi/\omega$, and not only the values of x_f and x_m must be compared, but also the values of $\dot{x}_f$ and $\dot{x}_m$.

It will be seen that the conditions to be imposed depend on the sign

of y_0 and the parity of m. The analysis may be conveniently divided into two cases: (1) $y_0 < 0$, m odd, which requires the weakest assumptions; (2) all the other possibilities. Suppose first that $y_0 < 0$ and m is odd. In this case both x_m and x_f start off positive from $t = 0$ and always agree in sign for values of t in the neighborhood of $t = km\pi/\omega$, $k = 0, 1, 2, \cdots$. As indicated in Fig. 2, $x_f + x_m$ is certainly positive

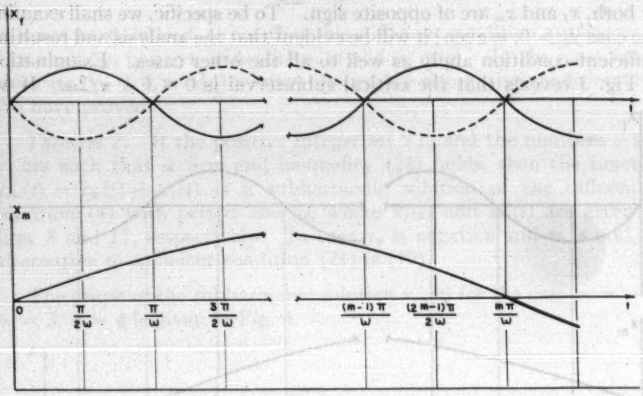

FIG. 2. Comparison of $x_m(t)$ and $x_f(t)$ for m odd. Solid and dashed curves in upper figure represent $x_f(t)$ in the cases $y_0 < 0$ and $y_0 > 0$, respectively.

for $0 < t < \pi/\omega$ and $(m-1)\pi/\omega < t < m\pi/\omega$; furthermore, if $x_f + x_m < 0$ anywhere in the interval $\pi/\omega < t < (m-1)\pi/\omega$, it must be in the subinterval $\pi/\omega < t < 3\pi/2\omega$, where x_m takes on its smallest values. Therefore $x_{sm} = x_m + x_f > 0$ throughout $0 < t < m\pi/\omega$ provided that $x_m(\pi/\omega) \geq \max |x_f|$, or, referring to Eqs. 8 and 17, that

$$\frac{\cos\dfrac{(m-2)\pi}{2\omega}}{\cos\dfrac{m\pi}{2\omega}} - 1 \geq \frac{y_0\omega^2}{1-\omega^2}. \tag{19}$$

Condition 19, then (together with $\omega > m$), is sufficient to insure the validity of the constructed subharmonic solution $x_{sm} = x_m + x_f$ in this case ($y_0 < 0$, m odd).

We now turn to the other possibilities ($y_0 < 0$, m even; $y_0 > 0$, m odd or even). For $y_0 > 0$ and m odd, $x_f(t)$ is represented by the dashed curve in Fig. 2; we see in this case that x_f and x_m disagree in sign both for t close to $t = 0$ and for t close to $t = m\pi/\omega$. In Fig. 3 the

graphs of $x_f(t)$ and $x_m(t)$ are compared for the case $y_0 > 0$, m even; the functions agree in sign for values of t close to $t = m\pi/\omega$, but disagree for t close to $t = 0$. It is not hard to see that for $y_0 < 0$ and m even the situation is reversed.

These three cases may be considered together because essentially they all present the same difficulty, which does not occur in the case $y_0 < 0$, m odd; namely, in the neighborhood of either $t = 0$ or $t = m\pi/\omega$ or both, x_f and x_m are of opposite sign. To be specific, we shall examine the case $y_0 > 0$, m even; it will be evident that the analysis and resulting sufficient condition apply as well to all the other cases. Examination of Fig. 3 reveals that the critical subinterval is $0 < t < \pi/2\omega$. If we

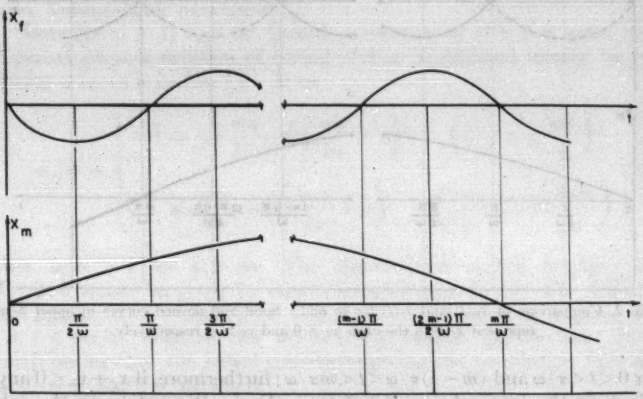

FIG. 3. Comparison of $x_m(t)$ and $x_f(t)$ in the case $y_0 > 0$, m even.

can insure $x_f + x_m > 0$, or $x_m > -x_f$, on this subinterval, the inequality will hold throughout $0 < t < m\pi/\omega$.

From the nature of the functions $x_f(t)$ and $x_m(t)$, we know that $-\dot{x}_f$ and $\dot{x}_m$ are both monotonic decreasing for $0 < t < \pi/2\omega$. Therefore, it is sufficient to specify that $\dot{x}_m(\pi/2\omega)$, the smallest value of $\dot{x}_m$ for $0 \leq t \leq \pi/2\omega$, be greater than or equal to $-\dot{x}_f(0)$, the largest value of $-\dot{x}_f$ in this interval; then $\dot{x}_m(t) > -\dot{x}_f(t)$ for $0 < t < \pi/2\omega$, and since $x_m(0) = x_f(0) = 0$, integration from 0 to t yields $x_m(t) > -x_f(t)$ for $0 < t < \pi/2\omega$, the desired result. Using Eqs. 8 and 17, we may write the sufficient condition in this case in the form

$$\dot{x}_m\left(\frac{\pi}{2\omega}\right) = \frac{\sin\dfrac{(m-1)\pi}{2\omega}}{\cos\dfrac{m\pi}{2\omega}} \geq -\dot{x}_f(0) = \frac{-y_0\omega^3}{1-\omega^2}. \qquad (20)$$

Since $\omega > m > 1$ and here $y_0 > 0$, this is a relation between positive quantities. In each case we must compare x_m and $|x_f|$; it follows then that Eq. 20 will apply to all three cases if we write it in the form

$$\frac{\sin \dfrac{(m-1)\pi}{2\omega}}{\cos \dfrac{m\pi}{2\omega}} \geqslant \frac{|y_0|\omega^2}{\omega^2 - 1}. \tag{21}$$

It is not hard, indeed, to see that this sufficient condition applies to all four cases: y_0 positive or negative, m odd or even. In the most favorable case ($y_0 < 0$, m odd) we may also use the alternative condition (19). We have proved

Theorem 2. If the positive integer $m(>1)$ and the numbers ω and y_0 are such that $\omega > m$ and inequality (21) holds, then the function $x_{sm}(t) = x_m(t) + x_f(t)$ is a subharmonic solution of the differential equation (4) with period $2m\pi/\omega$, where $x_f(t)$ and $x_m(t)$ are given by Eqs. 8 and 17, respectively. In case y_0 is negative and m is odd, an alternative to sufficient condition (21) is (19).

The graph of the subharmonic solution $x_{sm}(t)$ for the case $y_0 = -1$, $m = 3$, $\omega = 4$ is given in Fig. 4.

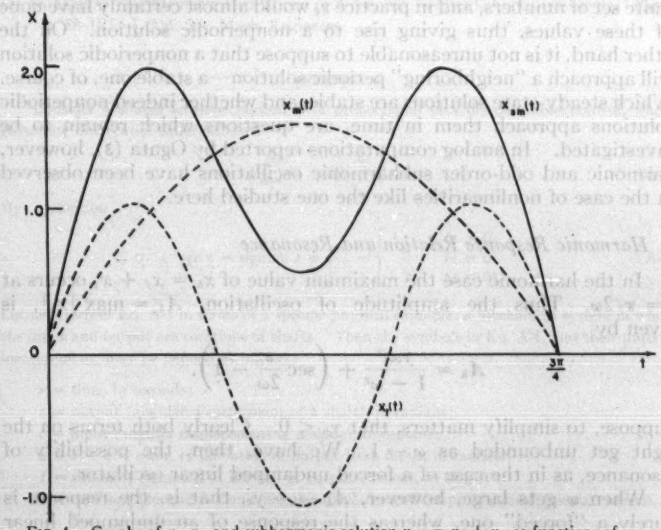

FIG. 4. Graphs of x_f, x_m and subharmonic solution $x_{sm} = x_f + x_m$ as functions of t when $y_0 = 1$, $\omega = 4$, $m = 3$.

III. CONCLUDING REMARKS

1. *Damping*

We have dealt here with a simplified mathematical model of a relay servomechanism; aside from the fact that the on-off element has been assumed to be ideal, damping has been neglected. It is intended, in a future investigation, to take damping into account.

2. *Stability*

For a given input function $y(t) = y_0 \sin \omega t$, Theorems 1 and 2 state that if ω lies between the integers M and $M + 1$ there exists a multiplicity of periodic solutions of Eq. 8, with the respective periods $2m\pi/\omega$, $m = 1, 2, \cdots, M$, provided that in each case y_0 satisfies an appropriate condition. Which of these steady-state solutions arises in a specific case (assuming that our undamped servomechanism can really exist) depends on the initial conditions $x(0)$, $\dot{x}(0)$—more precisely, on the value of

$$v_0 = \tan \frac{m\pi}{2\omega}.$$

These values of v_0 (for $m = 1, 2, \cdots, M$) comprise, however, a finite set of numbers, and in practice v_0 would almost certainly have none of these values, thus giving rise to a nonperiodic solution. On the other hand, it is not unreasonable to suppose that a nonperiodic solution will approach a "neighboring" periodic solution—a stable one, of course. Which steady-state solutions are stable, and whether indeed nonperiodic solutions approach them in time, are questions which remain to be investigated. In analog computations reported by Ogata (3), however, harmonic and odd-order subharmonic oscillations have been observed in the case of nonlinearities like the one studied here.

3. *Harmonic Response Relation and Resonance*

In the harmonic case the maximum value of $x_h = x_f + x_1$ occurs at $t = \pi/2\omega$. Thus the amplitude of oscillation, $A_h = \max |x_h|$, is given by

$$A_h = \frac{y_0\omega^2}{1 - \omega^2} + \left(\sec \frac{\pi}{2\omega} - 1\right).$$

Suppose, to simplify matters, that $y_0 < 0$. Clearly both terms on the right get unbounded as $\omega \to 1$. We have, then, the possibility of resonance, as in the case of a forced undamped linear oscillator.

When ω gets large, however, $A_h \to -y_0$, that is, the response is purely a "forced" one, whereas the response of an undamped linear system always contains forced and free components.

4. Subharmonic Response to High Frequencies

Suppose ω is much larger than 1, and $|y_0|$ is small. Then $\max|x_f|$ is small, and, as we have just seen, the harmonic response amplitude A_h is also small. It is possible, however, in this situation for the system to exhibit a subharmonic response of order m ($m = 1, 2, \cdots, M < \omega$). If m is rather large, $\max|x_m|$ is not negligible; then the amplitude, A_{sm}, of this subharmonic response $x_{sm} = x_m + x_f$ is approximately given by

$$A_{sm} \approx \max|x_m| = \sec\frac{m\pi}{2\omega} - 1. \qquad (\omega > m \gg 1)$$

As $m/\omega \to 1$, A_{sm} gets unbounded. This possibility of exaggerated, and perhaps destructive, response points up the practical importance of analyzing the various types of oscillations of which a nonlinear system is capable.

REFERENCES

(1) R. L. COSGRIFF, "Nonlinear Control Systems," New York, McGraw-Hill Book Co., Inc., 1958.

(2) J. C. WEST AND J. L. DOUCE, "The Mechanism of Sub-Harmonic Generation in a Feedback System," *Proc. I.E.E.*, Vol. 102, Pt. B, p. 569 (1955).

(3) K. OGATA, "Subharmonic Oscillations of Nonlinear Feedback Control Systems," Paper No. 58–IRD–7, Am. Soc. Mech. Engineers.

APPENDIX A

Consider a simple servomechanism with proportional plus discontinuous control, whose governing differential equation is

$$I\frac{d^2r}{d\tau^2} = -k^2e - T\,\mathrm{sgn}\,e. \qquad (A\text{-}1)$$

By definition

$$\mathrm{sgn}\,e = \mathrm{signum}\,e = \frac{e}{|e|} = \begin{cases} +1 & (e > 0) \\ 0 & (e = 0) \\ -1 & (e < 0) \end{cases} \qquad (A\text{-}2)$$

Let us interpret Eq. A-1 in terms of a specific physical example: a mechanical system in which the input and output are rotations of shafts. Then the symbols in Eq. A-1, and their units of measurement, may be defined as follows:

τ = time, in seconds;
r = output (angular displacement of a shaft), in radians;
c = input (angular displacement of a shaft), in radians;
$e = r - c$ = error, or deviation of output from input, in radians;
I = constant (moment of inertia) of output element, in slug-ft^2;
k^2 = constant (torque per unit angular displacement) of proportional control element, in lb-ft per radian;
T = constant (torque) of discontinuous control element in lb-ft.

The constants k^3, I and T are, of course, positive.

If we set $r = e + c$ in Eq. A-1, we obtain

$$I\frac{d^2e}{dt^2} + k^3e + T\,\mathrm{sgn}\,e = -I\frac{d^2c}{dt^2}. \tag{A-3}$$

By introducing the frequency parameter

$$\nu = \frac{k}{\sqrt{I}} \tag{A-4}$$

and the dimensionless quantities

$$t = \nu\tau, \qquad x = \frac{k^3e}{T}, \qquad y = \frac{k^3c}{T}, \qquad z = \frac{k^3r}{T}, \tag{A-5}$$

we may write Eq. A-3 in the simpler form

$$\ddot{x} + x + \mathrm{sgn}\,x = -\ddot{y}, \tag{A-6}$$

where $\cdot = d/dt$ and $\cdot\cdot = d^2/dt^2$. Sgn x = sgn e, that is, x and e are of the same sign, because k^3/T is positive. The variables t, x, y and z are referred to in the paper as the time, error, input and output, respectively.

APPENDIX B

We wish to find a condition on ω and c which insures $x_1(t) > -x_f(t)$ for $0 < t < \pi/\omega$ $(\omega > 1)$, where

$$x_1(t) = -1 + \cos t + \tan\frac{\pi}{2\omega}\sin t, \quad -x_f(t) = c\sin\omega t.$$

For convenience we have set

$$c = -\frac{y_0\omega^2}{1-\omega^2}.$$

Note that because $\omega > 1$, y_0 and c have the same sign.

First of all, since the graphs of $x_f(t)$ and $x_1(t)$ are symmetric about the line $t = \pi/2\omega$, we need only deal with the interval $0 < t < \pi/2\omega$. From examination of the graphs of $\sin\phi$ and $\cos\phi$ it is readily seen that

$$\frac{2}{\pi}\phi < \sin\phi < \phi \qquad \text{and} \qquad \cos\phi > -\frac{2}{\pi}\phi + 1$$

for $0 < \phi < \frac{\pi}{2}$, so that for $0 < t \leqslant \pi/2\omega < \pi/2$

$$x_1 > -1 - \frac{2}{\pi}t + 1 + \tan\left(\frac{\pi}{2\omega}\right)\frac{2}{\pi}t = -\frac{2}{\pi}t + \frac{2}{\pi}t\tan\frac{\pi}{2\omega}$$

and

$$-x_f < c\omega t.$$

It will be true, then, that $x_1 > -x_f$ for $0 < t < \pi/2\omega$ if

$$c\omega t \leqslant \frac{2}{\pi}\left(-1 + \tan\frac{\pi}{2\omega}\right)t$$

or

$$c \leqslant \frac{2}{\pi\omega}\left(-1 + \tan\frac{\pi}{2\omega}\right).$$

For $\omega < 2$ this upper bound on c (which corresponds to an upper bound on y_0) is positive. For $\omega > 2$ it is negative, but we know that we can always do better than this, for any negative c (or y_0) will do (see Part II, Sec. 3).

It is possible to improve this result by using sharper estimates of $\sin \phi$ and $\cos \phi$, obtained by integrating the inequalities given above. For example, if we integrate

$$\frac{2}{\pi}\phi < \sin \phi < \phi \qquad \left(0 < \phi < \frac{\pi}{2}\right)$$

from 0 to ϕ, we get

$$\frac{\phi^2}{\pi} < 1 - \cos \phi < \frac{\phi^2}{2}$$

or

$$1 - \frac{\phi^2}{2} < \cos \phi < 1 - \frac{\phi^2}{\pi} \qquad \left(0 < \phi < \frac{\pi}{2}\right).$$

The amplitudes of x_1 and x_f are $r - 1$ and $|c|$, respectively, and it is possible to show, indeed, that there is a value of ω, ω_0, such that for all $\omega > \omega_0 > 1$, the condition $c < r - 1$ implies $x_1(t) > -x_f(t)$ for $0 < t < \pi/\omega$.

CONTRIBUTION TO THE MODIFIED Z-TRANSFORM THEORY *

BY

E. I. JURY [1]

SUMMARY

In this discussion, a useful theorem applied to the modified z-transform (1, 2)[2] is introduced and verified. This theorem is fairly general and covers a few special cases that found many applications in the study of pulsed-systems.

These mentioned cases are discussed extensively and documented with examples and a short table in the Appendix. Various other cases of the theorem and its extension are pointed out in this paper.

CONVOLUTION THEOREMS

Theorem 1. Complex Convolution Theorem for the Modified Z-Transform[3]

$$\mathfrak{z}_m[f(t) \times h(t)] = \frac{1}{2\pi j}\int_\Gamma p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right)dp, \qquad 0 \le m \le 1, \quad z = e^{Ts}$$
$$= \frac{1}{2\pi j}\int_\Gamma p^{-1}H^*(p, m)F^*\left(\frac{z}{p}, m\right)dp \qquad (1)$$
$$f(t) = h(t) = 0 \quad \text{for} \quad t < 0$$

where Γ is a closed contour in a positive sense in the p-plane which encloses all the poles of $p^{-1}F^*(p, m)$, and $F^*(p, m) = \mathfrak{z}_m[f(t)]_{z=p}$, $H^*(z, m) = \mathfrak{z}_m[h(t)]$.

To prove the above theorem, we apply the complex convolution theorem to obtain the Laplace transform of multiplication of two time functions as follows (1–4):

$$\mathfrak{L}[f^*(t, m) \times h^*(t, m)] = \frac{T}{2\pi j}\int_{c-j(\pi/T)}^{c+j(\pi/T)} F^*(q, m)H^*(s - q, m)dq \qquad (2)$$

where the line integral $c - j\frac{\pi}{T}$, to $c + j\frac{\pi}{T}$, lies on an analytic strip in

* This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 18(600)–1521. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1] Department of Electrical Engineering, University of California, Berkeley, Calif.

[2] The boldface numbers in parentheses refer to the references appended to this paper.

[3] The complex convolution theorem for the z-transform case has been developed by Cypkin (2).

the complex q-plane that does not pass through or enclose the poles of $F^*(q, m)$ or of $H^*(s - q, m)$ and

$$F^*(q, m) = \mathfrak{L}[f^*(t, m)]|_{s=q} \tag{3}$$

$$H^*(s, m) = \mathfrak{L}[h^*(t, m)]. \tag{4}$$

The typical function $f^*(t, m)$ is defined as

$$f^*(t, m) = f[t - (1 - m)T] \times \delta_T(t)$$
$$= \sum_{n=0}^{\infty} f(nT - T + mT)\delta(t - nT) \qquad 0 \leq m \leq 1. \tag{5}$$

The Laplace transform of $f^*(t, m)$ and $h^*(t, m)$ is our definition of the modified z-transform of $f(t)$ and $h(t)$, respectively, thus

$$F^*(s, m) = \mathfrak{L}[f^*(t, m)] = \mathfrak{z}_m[f(t)],$$
$$\mathfrak{L}[h^*(t, m)] = \mathfrak{z}_m[h(t)] = H^*(s, m). \tag{6}$$

Because of the periodic nature of $\delta_T(t)$, the Laplace transform $F^*(s, m)$ is a rational function of e^{Ts}; thus, writing $F^*(s, m)$ in terms of e^{Ts}, replacing e^{Ts} by z we denote this function as $F^*(z, m)|_{z=e^{Ts}}$.

Changing the path of integration in Eq. 2 from the q-plane into the p-plane (1), (where $p = e^{Tq}$), and writing for $dq = \dfrac{de^{Tq}}{Te^{Tq}}$, we obtain:

$$\mathfrak{L}[f^*(t, m)h^*(t, m)] = \frac{1}{2\pi j}\int_\Gamma p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right) dp. \tag{7}$$

It should be noted that the line-integral in Eq. 2 is changed into a closed contour Γ in the p-plane as required by the change of the variable of integration. From the definition of the modified z-transform, the left side of Eq. 7 is equal to the $\mathfrak{z}_m[f(t)h(t)]$. Therefore:

$$\mathfrak{z}_m[f(t) \times h(t)] = \frac{1}{2\pi j}\int_\Gamma p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right) dp \tag{8}[4]$$

which proves the theorem.

For the case $h(t) = f(t)$,

$$\mathfrak{z}_m[f(t)^2] = \frac{1}{2\pi j}\int_\Gamma p^{-1}F^*(p, m)F^*\left(\frac{z}{p}, m\right) dp. \tag{9}$$

[4] When $h(t) = \mu(t)$, the left side of this equation reduces to $F^*(z, m)$.

Evaluation of the Complex Convolution Integral (1, 2)

To evaluate Eq. 8, we can make use of Cauchy's Integral Formula and evaluate the residues inside the circle which encloses only the poles of $p^{-1}F^*(p, m)$. An alternate procedure is to obtain the residues of $H^*\left(\frac{z}{p}, m\right)$ outside this circle, with path of integration taken in the negative sense, provided the integral is convergent. The choice between the two methods depends on the number of poles of each of the functions F^* and H^*.

Making use of the first procedure and assuming $F^*(p, m)$ has simple poles, then

$$\mathfrak{z}_m[f(t) \times h(t)] = \sum_{n=1}^{N} \frac{A(p_n, m)}{p_n B'(p_n)} H^*\left(\frac{z}{p_n}, m\right) \tag{10}[5]$$

where

$$F^*(p, m) = \frac{A(p, m)}{B(p)} \tag{11}$$

$$\begin{aligned} B(p) &= 0, p_1, p_2, \cdots p_n \\ B'(p_n) &= \left.\frac{dB}{dp}\right|_{p=p_n} \end{aligned} \tag{12}$$

In general for multiple poles cases, Eq. 8 becomes:

$$\mathfrak{z}_m[f(t) \times h(t)] = \sum_{\substack{\text{at poles of} \\ p^{-1}F^*(p, m)}} \text{Residue of } p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right) \tag{13}[5]$$

ILLUSTRATIVE EXAMPLES

Example 1

Assume

$$\begin{aligned} f(t) &= e^{-\alpha t} \\ h(t) &= e^{-\beta t} \end{aligned} \tag{14}$$

The modified z-transform of each of the above functions can be obtained by using the following relation or through the use of tables (1, 2).

$$\begin{aligned} F^*(z, m) &= \mathfrak{L}\left[f[t - (1 - m)T] \times \delta_T(t)\right]\Big|_{z=e^{rT}} \\ &= \sum_{n=0}^{\infty} f(n - 1 + m)Tz^{-n}, \quad 0 \leq m \leq 1 \end{aligned} \tag{15}$$

Thus from tables (1) or from Eq. 15,

$$F^*(z, m) = \mathfrak{z}_m[e^{-\alpha t}] = \frac{e^{-\alpha mT}}{z - e^{-\alpha T}} \tag{16}$$

[5] For these cases it is assumed that the residue of the pole at the origin is zero. If such a residue exists, then it should be modified to include its value.

$$H^*(z, m) = \mathfrak{z}_m[e^{-\beta t}] = \frac{e^{-\beta mT}}{z - e^{-\beta T}}. \tag{17}$$

From Eq. 8

$$\mathfrak{z}_m[e^{-\alpha t}e^{-\beta t}] = \frac{1}{2\pi j}\int_\Gamma \frac{e^{-\alpha mT}}{p - e^{-\alpha T}} \times \frac{e^{-\beta mT}}{\frac{z}{p} - e^{-\beta T}} p^{-1}dp. \tag{18}$$

Since $F^*(p, m)$ has a simple pole at $p = e^{-\alpha T}$, Eq. 10 can be applied which yields:

$$\mathfrak{z}_m[e^{-\alpha t}e^{-\beta t}] = \frac{e^{-(\alpha+\beta)mT}}{e^{-\alpha T}\frac{z}{e^{-\alpha T}} - e^{-\beta T}} = \frac{e^{-(\alpha+\beta)mT}}{z - e^{-(\alpha+\beta)T}}, \quad 0 < m < 1. \tag{19}$$

Example 2

Find the modified z-transform of the following function (2, 5)

$$\mathfrak{z}_m[tf(t)]$$

In this case (1):

$$\mathfrak{z}_m[t] = \frac{mTz - mT + T}{(z - 1)^2} \tag{20}$$

$$\mathfrak{z}_m[f(t)] = F^*(z, m). \tag{21}$$

From Eq. 8

$$\mathfrak{z}_m[tf(t)] = \frac{1}{2\pi j}\int_\Gamma \frac{mTp - mT + T}{(p - 1)^2}\frac{1}{p}F^*\left(\frac{z}{p}, m\right)dp. \tag{22}$$

The function of the integrand in the region of integration has a multiple[6] at $p = 1$, thus the sum of the residues equals (1),

$$\begin{aligned}\mathfrak{z}_m[tf(t)] &= \frac{\partial}{\partial p}\frac{mTp - mT + m}{p}F^*\left(\frac{z}{p}, m\right)\Big|_{p=1} \\ &= F^*\left(\frac{z}{p}, m\right)\frac{\partial}{\partial p}[mT - mTp^{-1} + Tp^{-1}]\Big|_{p=1} \\ &\quad + \frac{mTp - mT + T}{p}\frac{\partial}{\partial p}F^*\left(\frac{z}{p}, m\right)\Big|_{p=1} \\ &= F^*(z, m)[T(m - 1)] + m\frac{\partial}{\partial p}F^*\left(\frac{z}{p}, m\right)\Big|_{p=1}.\end{aligned} \tag{23}$$

[6] Since in most cases, $F^*\left(\frac{z}{p}, m\right)\Big|_{p=0} = 0$, it is assumed that the residue of the pole at the origin is zero.

It can be easily verified that for all forms of $F^*\left(\frac{z}{p}, m\right)\Big|_{p=1}$, the following is true

$$\frac{\partial}{\partial p} F^*\left(\frac{z}{p}, m\right)\Big|_{p=1} = -z\frac{\partial}{\partial z} F^*(z, m). \tag{24}$$

Therefore,

$$\mathfrak{z}_m[tf(t)] = T\left[(m-1)F^*(z, m) - z\frac{\partial}{\partial z} F^*(z, m)\right] = F_1^*(z, m). \tag{25}$$

Similarly for

$$\mathfrak{z}_m[t^2 f(t)] = T\left[(m-1)F_1^*(z, m) - z\frac{\partial}{\partial z} F_1^*(z, m)\right]. \tag{26}$$

By iterative extension of the above one can obtain, $\mathfrak{z}_m[t^k f(t)]$, k = positive integer.

APPLICATIONS OF THE THEOREM

Several special cases of wide interest can be obtained from the mentioned convolution theorem:

(*a*). *Application to obtain* $\mathfrak{z}[f(t)h(t)]$

The z-transform can be obtained from the modified z-transform by multiplying by z and letting $m = 0$ in $F_1^*(z, m)$.

$$F_1^*(z) = zF_1^*(z, m)\big|_{m=0} = z\mathfrak{z}_m[f(t)h(t)]\big|_{m=0}. \tag{27}$$

Alternately, in integral form $F_1^*(z)$ can be represented (2):

$$F_1^*(z) = \mathfrak{z}[f(t)h(t)] = \frac{1}{2\pi j}\int_\Gamma \frac{F^*(p)}{p} \times H^*\left(\frac{z}{p}\right) dp \tag{28}$$

where

$$F^*(p) = \mathfrak{z}[f(t)]\big|_{z=p} \tag{29}$$

$$H^*(z) = \mathfrak{z}[h(t)]. \tag{30}$$

(*b*). *Application to obtain* $\sum_{n=0}^{\infty} [f(nT)]^2$

Letting $h(t) = f(t)$ in Eq. 28, we obtain

$$\mathfrak{z}[f^2(t)] = \sum_{n=0}^{\infty} [f(nT)]^2 z^{-n} = \frac{1}{2\pi j}\int_\Gamma \frac{F^*(p)}{p} F^*\left(\frac{z}{p}\right) dp. \tag{31}$$

Letting $z = 1$ in the above equation to obtain

$$\sum_{n=0}^{\infty} [f(nT)]^2 = \frac{1}{2\pi j}\int_{\Gamma} F^*(p)F^*(p^{-1})p^{-1}dp, \tag{32}$$

where the contour-Γ is along the unit circle in the p-plane, because all poles of $p^{-1}F^*(p)$ should be inside the unit circle in order for the summation to exist.

The above equation is used extensively in evaluating the sum of the square error at the sampling instants in sampled-data control systems (2) and also in obtaining the noise power gain (6). The evaluation of the above integral (without obtaining the poles of $F^*(p)$) can be obtained as ratio of two determinants which has been tabulated in some publications (2, 5).

(c). *Application to obtain* (5).

$$\int_0^{\infty} f(t)h(t)dt = \frac{T}{2\pi j}\int_0^1\int_{\Gamma} F^*(p, m)H^*(p^{-1}, m)p^{-1}dpdm \tag{33}$$

or,

$$\int_0^{\infty} [f(t)]^2dt = \frac{T}{2\pi j}\int_0^1\int_{\Gamma} F^*(p, m)F^*(p^{-1}, m)p^{-1}dpdm \tag{34}$$

where the closed contour Γ is the unit circle in the p-plane.

The above theorem is analogous to Parseval's theorem for Fourier transforms. It has been applied for the analysis of integral-square error in sampled-data control systems (5).

To show the above application from the general theorem we first obtain the following equivalent form of the integral in the left side of Eq. 33.

$$\int_0^{\infty} f(t)h(t)dt = \sum_{n=0}^{\infty}\int_{nT}^{(n+1)T} f(t)h(t)dt$$

$$= T\sum_{n=0}^{\infty}\int_0^1 f(nT - T + mT)h(nT - T + mT)dm,$$

$$t = (n - 1 + m)T, \qquad n = \text{integer} \qquad 0 \le m \le 1 \tag{35}$$

On the other hand, the left side of Eq. 8 can be written as

$$\mathfrak{Z}_m[f(t)h(t)] = \sum_{n=0}^{\infty} f(nT - T + mT)h(nT - T + mT)z^{-n}. \tag{36}$$

If we let $z = 1$, multiply by T and integrate $\omega.r.$ to m we get:

$$T\int_0^1 \mathfrak{z}_m[f(t)h(t)]\big|_{z=1}dm = T\int_0^1 \sum_{n=0}^{\infty} f(nT - T + mT)h(nT - T + mT)dm. \quad (37)$$

Interchanging the order of summation and integration and using Eq. 8, we obtain

$$T\sum_{n=0}^{\infty}\int_0^1 f(nT - T + mT)h(nT - T + mT)dm = \frac{T}{2\pi j}\int_0^1\int_\Gamma F^*(p, m)H^*(p^{-1}, m)p^{-1}dpdm. \quad (38)$$

From Eq. 35, the above equals

$$\int_0^\infty f(t)h(t)dt = \frac{T}{2\pi j}\int_0^1\int_\Gamma F^*(p, m)H^*(p^{-1}, m)p^{-1}dpdm. \quad (39)^7$$

When $h(t) = f(t)$, the above yields Eq. 34. The evaluation of this integral is simplified if the order of integration in z and m is interchanged.[8] This is possible since the integrand is continuous along Γ and in the interval $0 \leq m \leq 1$. The evaluation of the first integral (5) is carried similar to Eq. 32.

Noting Example 2, for $\mathfrak{z}_m[tf(t)]$ and letting $h(t)$ in Eq. 39 be $[tf(t)]$, we get

$$\int_0^\infty t[f(t)]^2dt = \frac{T^2}{2\pi j}\int_0^1\int_\Gamma\left[(m-1)F^*(p, m) - p\frac{\partial}{\partial p}F^*(p, m)\right] \times F^*(p^{-1}, m)p^{-1}dpdm. \quad (40)$$

Similar relations can be obtained for higher time moments.

(*d*). *Application to obtain*

$$\mathfrak{z}_m\left[\int_0^t f(t)h(t)dt\right] = \frac{T}{2\pi j(z-1)}\int_0^1\int_\Gamma p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right)dpdm + \frac{T}{2\pi j}\int_0^m\int_\Gamma p^{-1}F^*(p, m)H^*\left(\frac{z}{p}, m\right)dpdm. \quad (41)$$

[7] As a special case when $h(t) = \mu(t) =$ unit step, we obtain

$$\int_0^\infty f(t)dt = T\int_0^1 F^*(p, m)\Big|_{p=1}dm, \quad \text{where} \quad F^*(p, m) = \mathfrak{z}_m[f(t)]\Big|_{z=p}.$$

[8] See Theorem 2 for evaluation of the integral $\omega.r.$ to m.

The above relationship can be obtained from writing the integral in a summation form as follows:

$$\int_0^t f(t)h(t)dt = T\left[\sum_{n=0}^{n}\int_0^1 f(n-1+m)Th(n-1+m)Tdm + \int_0^m f(n-1+m)Th(n-1+m)Tdm\right]. \quad (42)$$

The modified z-transform of the above yields[9]

$$\mathfrak{z}_m\left[\int_0^t f(t)h(t)dt\right] = \frac{T}{z-1}\int_0^1 \sum_{n=0}^{\infty} f(n-1+m)Th(n-1+m)Tz^{-n}dm + T\int_0^m \sum_{n=0}^{\infty} f(n-1+m)Th(n-1+m)Tz^{-n}dm \quad (43)$$

but the infinite summations are the modified z-transform of $f(t)\ h(t)$, thus:

$$\mathfrak{z}_m\left[\int_0^t f(t)h(t)dt\right] = \frac{T}{z-1}\int_0^1 F_1^*(z,m)dm + T\int_0^m F_1^*(z,m)dm \quad (44)$$

where

$$F_1^*(z,m) = \mathfrak{z}_m[f(t)h(t)] = \frac{1}{2\pi j}\int_\Gamma F^*(p,m)p^{-1}H^*\left(\frac{z}{p},m\right)dp. \quad (45)$$

For the special case when $h(t) = \mu(t)$ = unit step, the above yields

$$\mathfrak{z}_m\left[\int_0^t f(t)dt\right] = \frac{T}{z-1}\int_0^1 F^*(z,m)dm + T\int_0^m F^*(z,m)dm \quad (46)$$

where

$$F^*(z,m) = \mathfrak{z}_m[f(t)]. \quad (47)$$

Equation 46 also yields the modified z-transform (2) of $\frac{F(s)}{s}$. Extension to obtain $\mathfrak{z}_m\left[\frac{F(s)}{s^k}\right]$, by iterative process is straightforward.

(e) *Application to obtain*

$$F(s) = T\int_0^1 z^{-m+1}F^*(z,m)dm,$$

[9] Noting that (1), $\frac{1}{z-1}F^*(z) = \mathfrak{z}\left[\sum_{k=0}^{k=\infty-1} f(kT)\right]$.

where

$$F^*(z, m) = \mathfrak{z}_m[f(t)],\dagger \qquad F(s) = \mathfrak{L}[f(t)]. \tag{48}$$

To show the above, we multiply the left side of Eq. 33 by e^{-st} and integrate to obtain

$$\int_0^\infty f(t)h(t)e^{-st}dt$$

$$= T \sum_{n=0}^{\infty} e^{-nTs} \int_0^1 f(nT - T + mT)h(nT - T + mT)e^{(1-m)sT},$$

$$t = (n - 1 + m)T, \qquad 0 \leq m \leq 1 \tag{49}$$

or

$$\int_0^\infty f(t)h(t)e^{-st}dt$$

$$= T \int_0^1 z^{-m+1} \sum_{n=0}^{\infty} f(nT - T + mT)h(nT - T + mT)z^{-n}dm \tag{50}$$

but

$$\sum_{n=0}^{\infty} f(nT - T + mT)h(nT - T + mT)z^{-n} = \mathfrak{z}_m[f(t)h(t)],$$

therefore from Eq. 8 we get

$$F_1(s) = \int_0^\infty f(t)h(t)e^{-st}dt$$

$$= T \int_0^1 z^{-m+1} \int_\Gamma F^*(p, m)H^*\left(\frac{z}{p}, m\right) p^{-1}dpdm \tag{51}$$

or

$$F_1(s) = T \int_0^1 z^{-m+1}F_1^*(z, m)dm, \qquad 0 \leq m \leq 1 \tag{52}$$

where

$$F_1^*(z, m) = \mathfrak{z}_m[f(t)h(t)] \qquad \text{and} \qquad F_1(s) = \mathfrak{L}[f(t)h(t)]. \tag{53}$$

The above relation shows the equivalence between the Laplace transform and the modified z-transforms. The above has been shown derived differently by Cypkin (2). It is obtained here in a simple manner as a special case of the convolution theorem.

† $F^*(z, m)$ can be directly obtained from $F(s)$ by using the following convolution integral (1):

$$F^*(z, m) = \frac{z^{-1}}{2\pi j} \int_{c-j\infty}^{c+j\infty} F(p) \frac{e^{mpT}}{1 - e^{-T(s-p)}} dp \Big|_{z=e^{Ts}}.$$

Example

Knowing $F^*(z, m) = \dfrac{e^{-amT}}{z - e^{-aT}}$, obtain $F(s)$. From Eq. 52, $F(s)$ equals

$$F(s) = T\int_0^1 z^{-m}\frac{ze^{-amT}}{z - e^{-aT}}\,dm$$

$$= \frac{Ts}{z - e^{-aT}}\int_0^1 e^{-mT(a+s)\,dm}\Big|_{z=e^{Ts}} = \frac{1}{s+a}. \quad (54)$$

(*f*) *Application to obtain*

$$F(s) = \frac{T}{2\pi j}\int_{c-j(\pi/T)}^{c+j(\pi/T)} F^*(q)\frac{1}{s-q}\,dq, \quad \text{where} \quad F^*(q) = F^*(s)\big|_{s=q}. \quad (55)$$

Knowing $F^*(s) = \mathfrak{L}[f^*(t)]$ from the above equation, to obtain a continuous function $f(t)$, whose values coincide with $f^*(t)$ at the sampling instants only, the following condition should be satisfied:

$$f(t)u(t) = f^*(t)u(t), \qquad u(t) = \text{unit step} \quad (56)$$

or in Laplace transforms,

$$F(s) = \mathfrak{L}[f(t)u(t)] = \mathfrak{L}[f^*(t)u(t)]. \quad (57)$$

Using Eq. 2 to evaluate the right side of Eq. 57, we obtain (2)

$$F(s) = \frac{T}{2\pi j}\int_{c-j(\pi/T)}^{c+j(\pi/T)} F^*(q)\frac{1}{s-q}\,dq. \quad (58)$$

In evaluating the above equation only the poles in the linearized strip between $c - j\dfrac{\pi}{T}$ to $c + j\dfrac{\pi}{T}$ are considered.

Example

$$f^*(t) = e^{-\alpha nT}\delta_T(t), \qquad F^*(q) = F^*(s)\Big|_{s=q} = \frac{e^{Tq}}{e^{Tq} - e^{-aT}}. \quad (59)$$

Evaluating Eq. 58, for the poles of $F^*(q)$ we obtain

$$F(s) = T\frac{e^{Tq}}{\dfrac{d}{dq}\left[e^{Tq} - e^{-aT}\right]} \times \frac{1}{s-q}\Big|_{q=-\alpha} = \frac{1}{s+\alpha}. \quad (60)$$

Alternately evaluating for the pole $q = s$, we obtain

$$F(s) = T \times \frac{1}{T} \sum_{k=-\infty}^{k=\infty} \frac{1}{q + \alpha + jk\omega_r}\bigg|_{q=s,\ k=0} = \frac{1}{s + \alpha}, \qquad \omega_r = \frac{2\pi}{T}. \tag{61}$$

(g) *Application to obtain*

$$F^*(z, m) = \frac{1}{n} \sum_{k=0}^{n-1} F^*(z_n e^{j2\pi(k/n)}, m), \qquad z_n = e^{(T/n)s} = z^{(1/n)} \tag{62}$$

or

$$F^*(z) = \frac{1}{n} \sum_{k=0}^{n-1} F^*(z_n e^{j2\pi(k/n)}). \tag{63}$$

The above theorem signifies that, knowing the modified z-transform with respect to the period (T/n) (when n integer is larger than unity), one can obtain from it the modified z-transform of $F^*(z, m)$ with respect to the period T. This derivation is of importance in the analysis of multi-rate sampled-data feedback systems (1, 2).

The inverse of Eq. 62 indicates a time function which is formed by the product of the inverse of $F^*(z_n, m)$ and a sampled step-function with period T, that is, $f(k, m)T = f(k, m)T/n \times \mu(kT)$. Thus using the convolution integral we obtain:

$$F^*(z, m) = \frac{1}{2\pi j} \int_\Gamma F^*(p_n, m) p_n^{-1} \frac{1}{1 - (z_n/p_n)^{-n}}\, dp_n \tag{64}$$

where

$$F^*(p_n, m) = F^*(z_n, m)\big|_{z_n = p_n} = \mathfrak{L}[f^*(t, m)] = \mathfrak{z}_m[f(t)]\big|_{z_n = e^{(T/n)s}} \tag{65}$$

and

$$\frac{1}{1 - z^{-1}}\bigg|_{z = e^{Ts}} = \frac{1}{1 - z_n^{-n}}\bigg|_{z_n = e^{(T/n)s}} = \mathfrak{L}[\delta_T(t)] = \mathfrak{z}[u(t)]\bigg|_{z = z_n^n}. \tag{66}$$

Equation 64 can be written as:

$$F^*(z, m) = \frac{1}{2\pi j} \int_\Gamma F^*(p_n, m) p_n^{-1} \frac{-1}{(p_n/z_n)^n - 1}\, dp_n. \tag{67}$$

Integrating in a negative sense ω.r. to the poles of $\dfrac{1}{(p_n/z_n)^n - 1}$, which are

$$(p_n/z_n)^n = e^{j2\pi k}, \qquad k = 0, 1, 2, \cdots, n - 1,$$

we obtain the sum of the residues

$$= -\sum_{k=0}^{n-1} F^*(p_n, m)p_n^{-1} \frac{1}{\frac{d}{dp}(p_n/z_n)^n - 1}\Bigg|_{p_n = z_n e^{j2\pi(k/n)}} . \quad (68)$$

Thus,

$$F^*(z, m) = \frac{1}{n}\sum_{k=0}^{n-1} F^*(z_n e^{j2\pi(k/n)}, m). \quad (69)$$

For the case of z-transform, letting $m = 1$ in Eq. 69, we obtain

$$F^*(z) = \frac{1}{n}\sum_{k=0}^{n-1} F^*(z_n e^{j2\pi(k/n)}). \quad (70)$$

Further application of the theorem in the analysis of modulated discrete signals is of further use in the study of modulated sampled-data systems.

ADDITIONAL THEOREMS

Theorem 1a

$$\mathfrak{Z}_m\left[\frac{f(t)}{t}\right] = \frac{z^{m-1}}{T}\int_z^\infty z^{-m}F^*(z, m)dz, \qquad 0 < m < 1$$

$$\lim_{t\to 0} f(t) = 0, \qquad t = (n - 1 + m)T. \quad (71)$$

This theorem can also be derived from Eq. 8 as a special case when $h(t) = \frac{1}{t}$. However in using the convolution integral it is necessary to know $\mathfrak{Z}_m\left[\frac{1}{t}\right]$. To avoid this, the above is derived differently as follows:

$$F^*(z, m) = \sum_{n=0}^{\infty} f(nT - T + mT)z^{-n} = \sum_{n=1}^{\infty} f(nT - T + mT)z^{-n}, \quad (72)$$

since it is assumed that

$$f(t)\,|_{t=0} = 0.$$

Multiply the above equation by z^{-m} and integrate to obtain

$$\int z^{-m}F^*(z, m)dz = \int \sum_{n=0}^{\infty} f(nT - T + mT)z^{-(n+m)}$$

$$= \sum_{n=1}^{\infty} \frac{f(nT - T + mT)z^{-(n+m-1)}}{-n - m + 1}, \qquad 0 < m < 1 \quad (73)$$

or,

$$\frac{1}{T} z^{m-1} \int_z^\infty z^{-m} F^*(z, m) dz = \sum_{n=1}^{\infty} \frac{f(nT - T + mT) z^{-n}}{(nT - T + mT)}. \tag{74}$$

If $\dfrac{f(n-1+m)T}{(n-1+m)T} z^{-0}$ is added to both sides of Eq. 74, then this term has a contribution only at $m = 1$, which is the $\lim_{t \to 0} \dfrac{f(t)}{t}$, thus the right side of the above equation is $\mathfrak{z}_m \left[\dfrac{f(t)}{t} \right]$,

$$\mathfrak{z}_m \left[\frac{f(t)}{t} \right] = \frac{1}{T} z^{m-1} \int_z^\infty z^{-m} F^*(z, m) dz + \lim_{t \to 0} \frac{f(t)}{t}, \quad 0 < m \leq 1. \tag{75}$$[10]

For the case of z-transform we obtain, by substituting $m = 1$ in Eq. 75

$$\mathfrak{z} \left[\frac{f(t)}{t} \right] = \frac{1}{T} \int_z^\infty z^{-1} F^*(z) dz + \lim_{t \to 0} \frac{f(t)}{t},$$

where

$$F^*(z) = F^*(z, m)\big|_{m=1}. \tag{76}$$

Extension of Eq. 75 to obtain $\mathfrak{z}_m \left[\dfrac{f(t)}{t^k} \right]$ is evident (2).

Special cases which stem from this theorem are indicated in the following.

$$(a) \qquad \mathfrak{z} \left[\frac{f(t)}{t + hT} \right] = \frac{1}{T} z^h \int_z^\infty z^{-(1+h)} F^*(z) dz \tag{77}$$

where h is a positive constant larger than zero. The above is obtained from Eq. 75 by substituting for $m - 1 = h$, and using the z-transform case.

(b) When $z = 1$, relation (a) yields:

$$\sum_{n=0}^{\infty} \frac{f(nT)}{(n+h)T} = \frac{1}{T} \int_1^\infty z^{-(1+h)} F^*(z) dz, \qquad h > 0. \tag{78}$$

For the case $h = 0$, the above is modified to yield

$$\sum_{n=0}^{\infty} \frac{f(nT)}{nT} = \frac{1}{T} \int_1^\infty z^{-1} F^*(z) dz + \lim_{n \to 0} \frac{f(nT)}{(nT)}. \tag{79}$$

[10] When $m = 0$, we obtain $\mathfrak{z} \left(\dfrac{f(t-T)}{t-T} \right) = z^{-1} \mathfrak{z} \left(\dfrac{f(t)}{t} \right)$.

Example

Obtain, using Eq. 75,

$$\delta_m\left[\frac{1-e^{-\alpha t}}{t}\right], \quad f(t) = 1 - e^{-\alpha t}, \quad F^*(z, m) = \frac{1}{z-1} - \frac{e^{-\alpha m T}}{z - e^{-\alpha T}} \tag{80}$$

$$\int z^{-m} F^*(z, m)dz = \int \left\{\frac{z^{-m}}{z-1} - \frac{z^{-m}e^{-\alpha m T}}{z - e^{-\alpha T}}\right\} dz$$

$$= \int [1 + z^{-1} + z^{-2} + \cdots]z^{-m-1}dz$$

$$- \int [1 + z^{-1}e^{-\alpha T} + \cdots]z^{-m-1}e^{-\alpha m T}dz \tag{81}$$

for $0 < m < 1$

$$\int_z^\infty z^{-m} F^*(z, m)dz$$

$$= -\left[\frac{z^{-m}}{m} + \frac{e^{-\alpha T}z^{-m-1}}{m+1} + \frac{e^{-2\alpha T}z^{-m-2}}{m+2} + \cdots\right]e^{-\alpha m T}$$

$$+ \frac{z^{-m}}{m} + \frac{z^{-m-1}}{m+1} + \frac{z^{-m-2}}{m+2} + \cdots. \tag{82}$$

Finally,

$$\delta_m\left[\frac{1-e^{-\alpha T}}{t}\right] = \frac{z^{-1}}{mT}[1 - e^{-\alpha m T}] + \frac{z^{-2}}{(m+1)T}(1 - e^{-\alpha T}e^{-\alpha m T})$$

$$+ \frac{z^{-3}}{(m+2)T}(1 - e^{-2\alpha T}e^{-\alpha m T}) + \cdots +, \quad 0 < m \leq 1. \tag{83}$[11]$$

For the z-transform let $m = 1$ in Eq. 83, to obtain after closing the series

$$\delta\left[\frac{1-e^{-\alpha T}}{t}\right] = \alpha + \frac{1}{T}\ln\frac{z - e^{-\alpha T}}{z-1}. \tag{84}$$

A short table of z-transforms is given in the Appendix.

---

[11] It is of interest to note that when $\alpha \to \infty$, we obtain

$$\delta_m\left[\frac{1}{t}\right] = \frac{z^{-1}}{mT} + \frac{z^{-2}}{(m+1)T} + \frac{z^{-3}}{(m+2)T} + \cdots \frac{z^{-n}}{(m+n-1)T} + \delta_a(t) \quad \text{for} \quad 0 < m \leq 1.$$

For the case $m = 1$, we obtain

$$\delta\left[\frac{1}{t}\right] = \delta_a(t) + \frac{1}{T}\ln\frac{z}{z-1}, \quad \text{where} \quad \delta_a(t) = \lim_{\substack{\alpha \to \infty \\ t \to 0}}\left[\frac{1-e^{-\alpha t}}{t}\right].$$

In some publications (2) the z-transform of $\left[\frac{1}{t}\right]$ is defined only for $t > 0$, thus the infinite value of $\delta_a(t)$ is ignored.

Theorem 2

$$T\int_0^1 F^*(z, m)H^*(z^{-1}, m)dm = \mathfrak{z}[F(s)H(-s)], \quad 0 \le m \le 1. \quad (85)^{12}$$

The above theorem is an extension of the following theorem which is discussed in the literature (**1**, **5**, **7**, **8**).

$$T\int_0^1 F^*(z, m)F^*(z^{-1}, m)dm = \mathfrak{z}[F(s) \times F(-s)]. \quad (86)$$

Following similar steps for the derivation (**5**) of Eq. 86, one can easily obtain the theorem in Eq. 85.

The use of the above theorem is of advantage in the evaluation of the integral ω.r. to m in Eq. 39. Thus from Eqs. 39 and 85 we obtain

$$\int_0^\infty f(t)h(t)dt = \frac{T}{2\pi j}\int_\Gamma \int_0^1 F^*(p, m)H^*(p^{-1}, m)dm p^{-1}dp$$

$$= \frac{1}{2\pi j}\int_\Gamma \mathfrak{z}[F(s) \times H(-s)]|_{z=p}p^{-1}dp. \quad (87)$$

CONCLUSION

The complex convolution integral for both the z-transform and modified z-transform has been discussed in detail. The applications of the theorems for various fields such as sampled data systems, digital control systems, circuit theory and to the operational solution of difference equations both linear and non-linear, are possible using the techniques of this paper. It is of particular importance that the extension of the convolution integral for solving certain types of non-linear difference equations is warranted.

REFERENCES

(1) E. I. JURY, "Sampled-Data Control System," New York, John Wiley & Sons, Inc., 1958.
(2) JA. CYPKIN, "Theory of Pulse Systems," (in Russian), State Press for Physics and Mathematical Literature, Moscow, 1958.
(3) M. F. GARDNER AND J. L. BARNES, "Transients in Linear Systems," Vol. 1, New York, John Wiley & Sons, Inc., 1942.
(4) L. A. MACCOLL, "Fundamental Theory of Servomechanisms," Princeton, D. Van Nostrand Co., Inc., 1945.
(5) B. A. BHARUCHA, "Analysis of Integral-Square Error in Sampled-Data Control Systems," Electronics Research Laboratory Report, Series 60, Issue 206, University of California, Berkeley, June 1958.
(6) BARRY H. WHALEN, "Noise Power Gain of Discrete Filters," M. S. Thesis, University of California, June 1959.
(7) M. MORI, "Statistical Treatment of Sampled-Data Control Systems for Actual Random Inputs," *Trans. ASME*, Vol. 80, No. 2, February 1958.
(8) T. NISHIMURA, "Contribution to Statistical Design of Sampled-Data Control Systems," M.S. Thesis, University of California, Berkeley, 1958.

[12] $\mathfrak{z}[F(s)H(-s)]$ symbolizes the z-transform of a time function whose Laplace transform is $[F(s)H(-s)]$.

APPENDIX: A SHORT TABLE OF Z-TRANSFORMS OF CERTAIN FUNCTIONS

$f(t)$	$F^*(z, m)$	$F^*(z)$
$\frac{1-e^{-at}}{t}$	$\frac{z^{-1}}{mT}[1-e^{-amT}]$ $+\frac{z^{-2}}{(m+1)T}[1-e^{-aT}e^{-amT}]$ $+\frac{z^{-n}}{(m+n-1)T}[1-e^{-a(n-1)T}e^{-amT}]$ $+\cdots$	$a+\frac{1}{T}\ln\frac{z-e^{-aT}}{z-1}$, $a>0$
$\frac{1}{t}$, $t>0$	$\frac{z^{-1}}{mT}+\frac{z^{-2}}{(m+1)T}+\cdots\frac{z^{-n}}{(m+n-1)T}$ $=z^{m-1}\int_0^z\frac{p^{-m}}{1-p}\,dp$	$\frac{1}{T}\ln\frac{z}{z-1}$
$\frac{\sin at}{t}$	$\frac{1}{T}e^{-(1-m)Ts}\sum_{k=-\infty}^{k=\infty}\left[\tan^{-1}\frac{\alpha}{s+jk\omega_r}\right]e^{jmkT\omega_r}$ $z=e^{Ts}$, $\omega_r=\frac{2\pi}{T}$, $k=$integer	$a+\frac{1}{T}\left[\frac{\pi}{2}-\tan^{-1}\frac{z-\cos aT}{\sin aT}\right]$, $a>0$
$\frac{f(t)}{t}$	$\frac{1}{T}z^{m-1}\int_z^\infty z^{-m}F^*(z, m)dz+\lim_{t\to 0}\frac{f(t)}{t}$, $0\le m\le 1$	$\frac{1}{T}\int_z^\infty z^{-1}F^*(z)dz+\lim_{t\to 0}\frac{f(t)}{t}$
$\int_0^t f(t)dt$	$\frac{T}{z-1}\int_0^1 F^*(z, m)dm+T\int_0^m F^*(z, m)dm$	$\frac{Tz}{z-1}F^*(z)$
$t^k f(t)$	$T\left[(m-1)F_1^*(z, m)-z\frac{\partial}{\partial z}F_1^*(z, m)\right]$	$-Tz\frac{\partial}{\partial z}F_1^*(z)$
$k>0$, and integer	$F_1^*(z, m)=\mathfrak{Z}_m[t^{k-1}f(t)]$	$F_1^*(z)=\mathfrak{Z}[t^{k-1}f(t)]$

THE FIRST ELECTRIC COMPUTER, A MAGNETOLOGICAL ANALYSIS *

BY

GEORGE W. PATTERSON [1]

SUMMARY

Mechanisms designed to carry out the ordinary operations of decimal arithmetic were invented several centuries ago, and indeed Charles Babbage, in England, conceived automatic sequence control over one hundred years ago. The modern electronic high-speed computer based on binary Boolean logic has other ancestors besides decimal mechanical calculators in its pedigree, however. The mechanical realization of binary logic is generally credited to William Stanley Jevons, British economist and logician, in 1869.[2] This mechanism is an ancestor of the modern computer on the binary side of the family. It has recently come to light that an American university professor designed what is probably the first electric binary logic machine, and this is very likely the first design of an electric data processor of any kind.

Professor Alonzo Church of Princeton University has discovered a cardboard schematic diagram among the papers and manuscripts of Allan Marquand (1853–1924), former Professor of Archeology and the History of Arts at Princeton University. This diagram, believed to have been prepared about 1885, was described and reproduced in facsimile in a communication to the magazine *Science*.[3] Marquand had constructed an earlier mechanical logic machine[4] that still is extant. It seems quite plausible that the electrical design was based on the earlier mechanical one, but it is not clear whether the electrical machine was ever built, or, if built, whether it operated successfully. The present writer would give the designer the benefit of the doubt, and believes that the proper values of the parameters (not given on the schematic), if not known initially, would be soon discovered.

The explanation of the mode of operation of the machine is not straightforward since a number of assumptions must be made concerning the magnetomotive forces operating on the relay armatures. The basic difficulties arise from two facts: (*a*) The truth-functional operations are not mechanized in terms of controlled-contact networks, but

* This investigation was performed under the sponsorship of the Department of the Army, U. S. Army Signal Research and Development Laboratories, Contract No. DA36-039 SC-72344.

[1] The Moore School of Electrical Engineering, The University of Pennsylvania, Philadelphia, Pa.

[2] MARTIN GARDNER, "Logic Machines and Diagrams," New York, McGraw-Hill Book Co., Inc., 1958, p. 91.

[3] W. MAYS, "The First Circuit for an Electrical Logic Machine," *Science*, Vol. 118, pp. 281–282 (1953).

[4] ALLAN MARQUAND, "A New Logical Machine," *Proc. Am. Acad. Arts and Sci.*, Vol. 21, pp. 303–307 (1885–6). See also Gardner, *op. cit.*, pp. 106–113.

in terms of threshold elements whose logical properties are described in a recent paper[5]; (*b*) the relays are not exclusively truth functional in their operation, but also realize storage (memory) devices. These two points will be considered in more detail. Concerning (*a*), the description of a threshold element is as follows

$$h \equiv \{\sum_{i=1}^{I} \mathrm{c}(f_i) + \sum_{j=1}^{J} c(\bar{g}_j) \geqslant T\}$$

In this formula, h names the proposition:

the output is in its fiducial state,

while f_i and g_j correspond, respectively, to propositions of the form:

the i^{th} excitatory input is in its fiducial state,

the j^{th} inhibitory input is in its fiducial state.

By fiducial state, we mean that a signal is present. The integral constants, I, J, T correspond, respectively, to the total number of excitatory inputs, the total number of inhibitory inputs, and the threshold. The bar or vinculum over the g_j serves to negate the proposition, that is to say, $\bar{g}_j$ will be true if and only if g_j is false, that is, the j^{th} inhibitory input is not in its fiducial state, that is, it is in the antifiducial (no signal) state. The expressions $\mathrm{c}(f_i)$ and $c(g_j)$ denote characteristic functions of the propositions. $\mathrm{c}(f_i)$ is 1 if f_i is true and is 0 if f_i is false. From what has just been said concerning the bar, it follows that $\mathrm{c}(\bar{g}_j)$ is 1 if g_j is false and is 0 if g_j is true.

It turns out then that $\sum_{i=1}^{I} \mathrm{c}(f_i)$ is simply the total number of excitatory inputs out of I that are in, and $\sum_{j=1}^{J} \mathrm{c}(\bar{g}_j)$ is the total number of inhibitory inputs out of J that are not in, their fiducial states.

T is the threshold. If the total number of inputs that are either excitatory and in, or inhibitory and not in, their fiducial states is greater than or equal to T, then and only then is h true. For example, suppose T is 5, I is 4 and J is 2. Then if all four of the excitatory inputs are energized, and one or both of the inhibitory inputs are not energized the output will be energized.

In the present design, the inputs are physically realized by the five windings on the relay core. According to the diagram in Fig. 1, these windings all contribute ampere-turns of the same sign. This assumes that, and there is no reason to the contrary, the diagram accurately reflects coil and battery polarity. In these circumstances it is convenient to assume, for each literal winding, that its fiducial state corresponds to current flowing in it, and that the fiducial state of the relay is the energized state (armature closed against core). Then all the

[5] A. W. BURKS AND I. M. COPI, "The Logical Design of an Idealized General-Purpose Computer," JOUR. FRANKLIN INST., Vol. 261, pp. 299–314 and 421–436 (1956).

literal inputs are excitatory, $J = 0$, and the relay realizes a symmetric truth function, one that is true provided T or more inputs are energized and false otherwise.

Threshold elements of this simple type can physically realize, among other things, generalized n-input "or" (alternation or disjunction, with $T = 1$) and generalized n-input "and" (conjunction, with

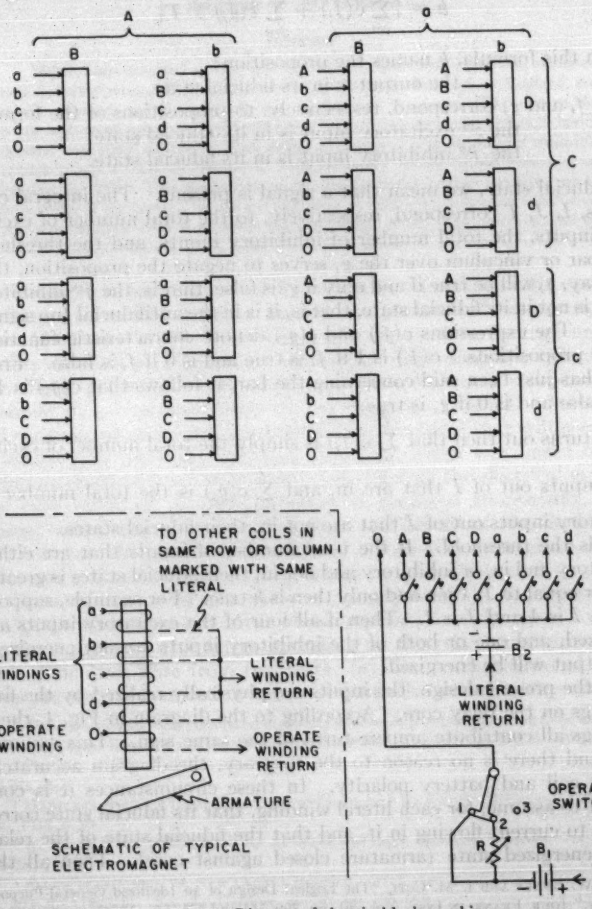

FIG. 1. Diagram of electrical logic machine.

$T = n$), but cannot realize negation. For this reason it is necessary to provide the machine with a pair of inputs for each variable. If the variable is affirmed, one key is depressed, if denied, the other key is depressed. There is no physical bar to depressing both keys simultaneously, leading to a tautologous (or contradictory) term. If neither key is depressed the term does not contain the variable. This is an advantage since it permits entering a formula in a normal form that is not canonical (Boolean, fully developed about all variables, distinguished).

It is clear that different windings may contribute different amounts of magnetomotive force. It seems likely that all windings connected to literal keys produce the same number of ampere-turns, say l. We use this as a unit of magnetomotive force in the sequel. The fifth winding on each core is connected to the operate switch, and is differently arranged. The switch produces two distinct non-zero values of mmf (contacts 1 and 2 in Fig. 1), as well as zero mmf (contact 3).

What is the relative effectiveness (O_1, O_2) of these positions compared to the other four literal windings? A natural assumption is that position 1 is twice as effective as a single literal winding, ($O_1 = 2l$) since battery B_1, which is connected to contact 1, appears to have twice as many cells as battery B_2, which is connected to the literal keys; for simplicity assume that $O_2 = l$. These assumptions are believed unworkable for reasons yet to be explained. Consider the magnitude of the threshold T in the light of (b) above: the relays are not exclusively truth functional in their operation, but also realize storage (memory) devices. The input necessary to cause a relay to close is greater than that required to retain it in the closed condition. If T is the operate threshold, and the imminent release point is t, we may have T/t as large as 10. Thus if T or more is required to cause the relay to pull in, we cannot state that less than T will cause the relay to drop out. Excluded middle is violated and the state of the relay for inputs between t and T is indeterminate and depends on the past, and this permits data storage or memory. The variable airgap in the magnetic circuit produces a double-valued function of flux *versus* mmf, and more important, a double-valued function of armature position *versus* mmf. This is analogous to the hysteresis loop in a rigid magnetic structure and the double-valuedness is what makes the memory or storage function possible; the quiescent point is located in the double-valued region.

The value of the ratio T/t must be established. According to Mays, "Position 2 sustains the pointers at the horizontal . . . but does not raise the fallen pointers." Also, "Position 1 raises all the fallen pointers to the horizontal." The assumptions are acceptable and yield $t \leqslant O_2 < T \leqslant O_1$. If we retain the first assumptions concerning the values of 0_1 and 0_2, we obtain $t \leqslant l < T \leqslant 2l$, and simple assumptions would be to set $t = \frac{l}{2}$, $T = \frac{3l}{2}$ to give ample margins of operation. These

assumptions yield $T/t = 3$, which is conservative. Unfortunately, no method of successful operation has been discovered for this set of assumptions. Now let us suppose that $t \leqslant O_2 = l < T \leqslant O_1$ as before but let $t \leqslant l < 5l < T$. This gives $5 < T/t$, which is physically still feasible. To give wide operating margins we might set $t = 0.5$, $T = 5.5$ with $T/t = 11$, although this is a rather extreme figure.

The machine is conveniently viewed as a truth-table schema consisting of sixteen cells, or atoms, upon which we wish to impress an arbitrary four-variable truth function expressed in (not necessarily distinguished) conjunctive normal form. Each conjunct is the disjunction of eight or fewer literals (the variables, negated and unnegated). Usually a variable will not appear both negated and unnegated in a disjunction, and as stated before one of the limitations of simple threshold elements of the type just described is that they cannot realize negation directly. Hence the absence of an A signal does not imply the presence of the a (that is, a not-A) signal. This also has the advantage that conjuncts may omit variables entirely, and this would not be possible if we had four two-position switches for the variables, instead of eight one-position switches.

We successively impress disjunctions of literals on the machine, seriatim, yielding a truth table that corresponds to a conjunction of the new disjunction with the result obtained from the previous steps. If we start with a tautology (corresponding to a conjunction of zero conjuncts) then the accumulated result will represent the truth table obtained by conjoining all the disjunctions with a tautology. This process must be monotone in the sense that each conjunct can change a relay indication from energized (truth) to de-energized (falsity) but not vice-versa. The explanation is recursive. The starting configuration is obtained by putting the operate switch on 1. This produces on all coils an mmf O_1 such that $T \leqslant O_1$ and all sixteen relays operate and indicate truth. The operate switch is then instantaneously (make-before-break contact) placed on 2, yielding an mmf O_2 such that $t \leqslant O_2$, and all relays continue to indicate truth unless O_2 is removed. This manipulation impresses a tautology (corresponding to a conjunction of zero conjuncts) on the machine, with all relays indicating truth.

Now consider an arbitrary truth table on the machine with operate switch at position 2 (mmf O_2); some of the relays indicate truth and some falsity. We depress some (non-void) combination of the literal keys. (Note that coils on the A relays are energized by the a key, etc.) This applies an additional magnetomotive force of at least l on each coil that corresponds to one or more literals. In other words, a force of l (or greater) is applied to all atoms that imply the disjunction, and to no others. An atom indicating falsity (de-energized relay) will not be changed to indicate truth, however, since the mmf is at most $5l < T$ and the relay remains de-energized. The operate switch is now set to 3.

All the atoms that do not imply the disjunction are de-energized and indicate falsity. In this way the only atoms remaining energized and indicating truth are those corresponding to the conjunction of the previous result with the current disjunction. (These disjunctions are called premisses by Mays and also by Marquand himself, who views them as negations of conjunctions.) It remains only to store the new truth table. This is accomplished by returning the operate switch to position 2. The magnetomotive force is the same as before but with the significant difference that zero or more relays have released. We may now release the literal keys, and the storing mmf O_2 retains the result until another disjunction is conjoined. If for any reason we fail to retain any literal keys in the depressed state while setting the operate switch to zero, all cells go to zero as might be expected since we are conjoining a void disjunction, and the latter corresponds to a contradiction.

A comparison will reveal that this analysis of the behavior of the machine is different from that described in Mays's paper. No engineering assumptions consistent with his description and at the same time yielding a logically workable machine have been discovered. In order to discuss Mays's description we must translate his vocabulary (consistent with Marquand's) into our own. We use disjunction (a conjunct in the conjunctive normal form) for Mays's *premiss*.

Marquand considers the premisses to be denials of conjuncts. He expresses each premiss in the form

$$\mathrm{o}(A, B, C, D) \Rightarrow F,$$

where o is some truth-functional operator in four or fewer variables equivalent to a conjunction of literals. The premiss itself is equivalent to $\sim \mathrm{o}(A, B, C, D)$.

In Mays's analysis of the machine the operate switch is first set to 1, and then the individual letters of the first premiss are depressed seriatim. These depressions are interlaced with pairs of manipulations of the operate switch, first to 3 and then to 2.

Suppose the first premiss is $A \Rightarrow B$, which is equivalent to $\sim(A\bar{b})$. In our terminology this is the first disjunction, $a \vee B$, of a conjunctive normal form that may contain other conjuncts. In general, we may assume that the machine has an arbitrary truth-functional configuration already impressed upon it, say $\mathrm{o}^*(A, B, C, D)$, and the operate switch is on 2. With the operate switch resting on contact 2 those and only those relays consistent with $\mathrm{o}^*(A, B, C, D)$ are operated, although all are drawing the same current since no literal keys are depressed. According to Mays, the premiss is recorded on the machine in the following way. First the A key is depressed. This sends additional current thru the a coils (but not enough to pick up any of the relays).[6]

[6] The parenthetical remarks are applicable to the present analysis of the mode of operation of the machine.

The operate switch is then set to position 3, and only the a coils now receive any current. (All non-a relays that had previously been operated drop out, leaving only $o^*(A, B, C, D)$ & a impressed on the machine). The operate switch is then set to position 2. (This restores the holding current, storing $o(A, B, C, D)$ & a, and the A key may now be released, since it cannot ever pick up any relays.) Literal key b is now depressed and this sends extra current thru the B coils (but not enough to pick up any of the relays). The operate switch is again set to position 3, and only the B coils now receive any current. (Any non-B relays that had previously been operated drop out, leaving only $o(A, B, C, D)$ & a & B impressed on the machine.) At this point Mays states "the Ab pointers fall to the vertical." I assume this is intended to mean, when $o(A, B, C, D)$ is a tautology, that is, the initially cleared state of the machine, that *only* the Ab pointers fall to the vertical, that is to say, all the other relays remain energized. Under the present interpretation the AB and the Ab pointers (that is, the A pointers) fall on the first step and the ab pointers fall on the second step. This leaves *only* the aB pointers still horizontal.

If the premise $a \vee B$ is to be recorded, we must not drop either the AB or ab pointers. Under the present interpretation this can be done in one step. With $o(A, B, C, D)$ stored in the machine, we have the operate switch resting on 2. Literal keys A and b are *simultaneously* depressed. This sends additional current through the a coils and also through the B coils (but not enough to pick up any of the relays, even the aB relays, which get a double increment of current). The operate switch is then set to position 3, and the only coils receiving current now are the a coils and the B coils. All relays corresponding to atoms that imply $a \vee B$ ($\equiv \sim (Ab)$) that had previously dropped out remain de-energized, and relays corresponding to atoms that do not imply $a \vee B$ drop out now leaving

$$o(A, B, C, D) \ \& \ (a \vee B)$$

impressed on the machine. The operate switch is then set to position 2 and the above result is stored.

The principal objection to Mays's analysis is that there seems to be no distinction between the & and $\vee$ operators. For instance, suppose we wish to impress $\sim(Ab)$ & $\sim (BcD)$ which is equivalent to $(a \vee B)$ & $(b \vee C \vee d)$. In the present analysis the A and b keys are depressed simultaneously, the operate key manipulated, the A and b keys released, then the B and c and D keys are depressed simultaneously; in other words each premiss is handled as a unit. On the other hand, Mays says that each literal is recorded separately, even within a premiss. When it comes to starting a new premiss there is no clear specification of how to start. He says "We must thus continue to impress as many of the premisses of a problem as we please. . . ." But he has just been

talking about individual literals of a premiss, and if we continue in the same way we are just augmenting the first premiss with more literals. Furthermore, since he manipulates the operation switch between literals, he is, under the present physical assumptions, considering each literal to be a separate premiss. This means that $\sim(Ab)$ instead of being entered as $a \vee B$, is entered as $a \,\&\, B$, that is, as $\sim A \,\&\, \sim b$, and the DeMorgan law is violated.

In conclusion certain simplifications and generalizations of the machine may be described.

1. The battery B_1 is dispensable if we substitute a mechanical device to close the relay armatures, each of which will remain closed with the operate switch on 2. Return all coils and contact 2 to battery B_2.

2. Neither Marquand nor Mays mentions that, owing to duality, the machine can be used to express truth functions in (not necessarily distinguished) disjunctive as well as conjunctive normal form. Let operated relays correspond to falsity (redesignating the fiducial state), capital letters to true variables, simultaneous depression of literal keys to conjunction, and the use of the operate key to disjunction. The effect now is to accumulate true atoms, by losing falsities as the relays drop out. This is a more natural mode of operation, if we wish to compare the machine with present-day logic machines.

3. The machine generali zesnaturally to n variables, with the limitation that the amount of equipment approximately doubles when n goes to $n + 1$.[7] Also, $T/t > n + 1$.

[7] This limitation is not inherent. See WILLIAM MIEHLE, "Burroughs Truth-Function Evaluator," *J. Assn. Comp. Machinery*, Vol. 4, pp. 189–192 (1957), and ARTHUR W. BURKS, DON W. WARREN, JESSE B. WRIGHT, "An Analysis of a Logical Machine Using Parenthesis-Free Notation," *Math. Tables Aid. Comp.*, Vol. 8, pp. 53–57 (April 1954).

BARTOL RESEARCH FOUNDATION
MARTIN A. POMERANTZ, DIRECTOR

LIFETIME OF THE 161-KEV EXCITED STATE IN Sn^{117} *

BY

F. R. METZGER

ABSTRACT

Using the thermal resonance fluorescence method, the mean life of the 161-kev M1 transition in Sn^{117} has been measured as $(5.7^{+1.4}_{-1.4}) \times 10^{-10}$ second. In arriving at this value, a total conversion coefficient $\beta_T = 0.15$ was assumed.

INTRODUCTION

The 161-kev transition in Sn^{117} is known to be almost pure M1, the reported (1)[1] mixing ratio being $E2/M1 = 0.0015^{+0.0015}_{-0.0010}$. It is, therefore, not surprising that Fagg *et al.* (2), using 4-Mev alpha particles, were unable to observe Coulomb excitation of the 161-kev level in spite of the use of enriched Sn^{117}. Since no information concerning the lifetime of this *l*-forbidden M1 transition was available, a measurement using the resonance fluorescence technique was undertaken. Both the thermal and the centrifuge methods are suitable for this isotope. For the available source strength and source purity, the thermal method was judged to be capable of giving a more accurate result for the mean life.

The angular distribution of the resonance radiation for a spin sequence $\frac{1}{2} - \frac{3}{2} - \frac{1}{2}$ is of the form $W(\theta) = 1 + A_2 P_2 (\cos\theta)$, that is, does not contain higher terms for any E2/M1 mixture. The total scattering cross section may, therefore, be obtained from a single measurement at a scattering angle ($\theta \sim 125°$) for which P_2 vanishes.

EXPERIMENTAL PROCEDURE

A. General

When the same isotope is used as the source of the exciting radiation and as the scattering material, the resonance condition is usually, at least partially, destroyed by the recoil energy losses in the emission and absorption processes. The thermal method (3) is one way of restoring the resonance condition. In analyzing the results obtained with the thermal method one assumes that the temperature dependence of the overlap of emission and absorption line is that calculated on the basis of Maxwellian velocity distributions in source and scatterer. A measure of this overlap is the average cross section (averaged over the

* Assisted by the U. S. Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1] The boldface numbers in parentheses refer to the references appended to this paper.

incident line) which may be written in the form (3)

$$\sigma_{Av} = \frac{g_2}{g_1}\frac{\lambda^2}{4\pi^{1/2}}\cdot\frac{\Gamma_0^2}{\Gamma}\cdot\frac{1}{E_\gamma}\left(\frac{Mc^2}{2k(T_1+T_2)}\right)^{1/2}\exp\left(-\frac{E_\gamma^2/Mc^2}{2k(T_1+T_2)}\right) \quad (1)$$

where g_2 and g_1 are the statistical weights of excited state and ground state, respectively, λ is the wavelength of the gamma radiation of energy E_γ, Γ is the total width of the level, Γ_0 the partial gamma width of the transition to the ground state, M the mass of the recoiling nucleus, k Boltzmann's constant, c the velocity of light, T_1 the absolute temperature of the source, T_2 the absolute temperature (4) of the scatterer.

The presence of non-resonant elastic scattering, of room background and of contributions from other transitions makes it difficult to determine σ_{Av} on the basis of a single measurement. It is much easier to measure the change in σ_{Av} brought about by a change in the source temperature and to evaluate Γ_0^2/Γ from this change of σ_{Av}.

B. Source Preparation

Ten milligrams of SnO_2, enriched[2] to 92.62 per cent in Sn^{116}, were irradiated for a period of one week in the Oak Ridge National Laboratory Research Reactor in a slow neutron flux of 1.7×10^{14} n/cm^2 sec. In this way approximately 0.5 millicuries of 14-day Sn^{117} were produced. Impurities giving rise to gamma rays of higher energies were found to represent less than one per cent of the initial activity. They did, however, through Compton scattering, contribute appreciably to the non-resonant background counting rate in the pulse height region corresponding to the 161-kev radiation. For the studies reported here, no effort was made to remove these impurities, especially since an appreciable fraction of the hard gamma radiation originated from 118-day Sn^{113}.

C. Geometry

The experimental arrangement used for the scattering studies is shown in Fig. 1. The cylindrical tin scatterer was 8.3 in. in diameter, 2 in. high, and 0.080 in. thick. Cd was used for the comparison scatterer. Measurements were taken with the source at room temperature and with the source at 700°C. The counting rates were of the order of 1000 counts per minute in a channel comprising approximately the full width at half maximum of the 161-kev line. The increase in the scattering from the tin scatterer upon heating of the source from room temperature to 700°C. amounted to approximately one per cent of the total counting rate. The cross sections σ_{Av}(700°C.) and σ_{Av}(27°C.) were

[2] Obtained from the Stable Isotopes Division of the AEC.

determined absolutely, that is, not by comparison with the Rayleigh cross section. The absolute determination was chosen because it was felt that the identification of the non-resonant counting rate in the elastic peak with the Rayleigh cross section is highly problematical even for a source emitting a single gamma ray. If, as in the case of the Sn^{117} source used in this experiment, higher energy gamma rays are present, this identification would clearly be incorrect. In order to obtain the

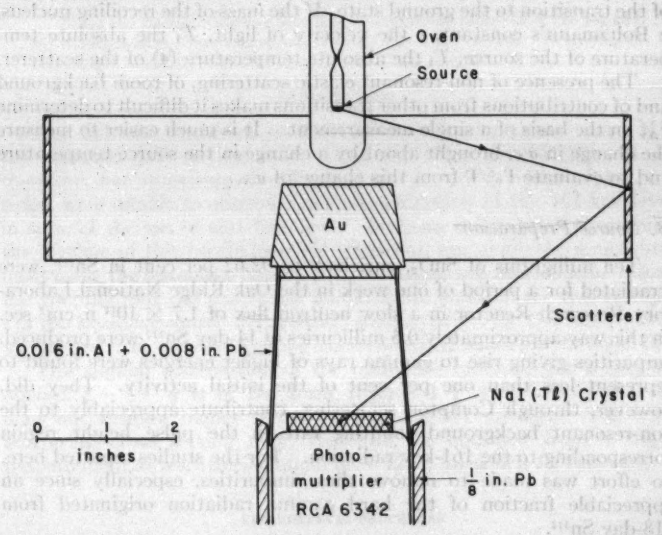

FIG. 1. Experimental arrangement. The temperature of the source was measured with a Chromel P-Alumel thermocouple.

absolute cross sections, the counting rate with the detector, a $\frac{1}{4}$ in. thick, $1\frac{1}{2}$ in. diameter NaI(Tl) crystal, exposed to the direct beam from the unshielded source was compared with the resonance counting rate, taking into account the geometry of Fig. 1. Since the exciting radiation and the resonantly scattered gamma rays are essentially of the same energy, this comparison does not involve the knowledge of the efficiency of the detector for different gamma ray energies.

RESULTS AND DISCUSSION

With $T_1 + T_2 = 610°$ for the experiment carried out at room temperature, and $T_1 + T_2 = 1281°$ for the high temperature experiment,

one calculates from Eq. 1:

$$\Delta\sigma_{Av} = \sigma_{Av}(700°C.) - \sigma_{Av}(27°C.) = 1.38 \times 10^{-19}\Gamma_0^2/\Gamma \text{ cm}^2 \quad (2)$$

where Γ_0 and Γ are expressed in ev.

For the geometry depicted in Fig. 1, the resonance counting rate difference is calculated as

$$\Delta N_{RES} = 1.95 \times 10^{19} N_0 A \Delta\sigma_{Av} \quad (3)$$

where $N_0 = C(d) \times d^2$ is the product of the direct beam counting rate and the square of the distance source-detector (in inches) employed in the direct beam measurement, and A is the natural abundance (0.0761) of Sn^{117}. Before calculating N_0 from the direct beam counting rate $C(d)$, the measured $C(d)$ was reduced by 3 per cent. This reduction represents the estimated contribution of the gamma rays from the highly converted 159-kev M4 transition preceding the 161-kev M1 radiation of interest. In arriving at this contribution, theoretical conversion coefficients (5) for pure M4 and pure M1 transitions were used. Combining Eqs. 2 and 3, one has

$$\Gamma_0^2/\Gamma = 4.88\Delta N_{RES}/N_0. \quad (4)$$

With the theoretical (5) total conversion coefficient for a pure M1 transition, $\beta_T = 0.15$, Γ_0/Γ takes on the value $\Gamma_0/\Gamma = 0.87$, and Eq. 4 goes over into

$$\Gamma_0 = 5.61\Delta N_{RES}/N_0. \quad (5)$$

From ten separate runs, an average value of $(20.6 \pm 7.7) \times 10^{-8}$ was obtained for the ratio $\Delta N_{RES}/N_0$. With Eq. 5 this then leads to a partial width $\Gamma_0 = (1.16 \pm 0.43) \times 10^{-6}$ ev, or a mean life

$$\tau_\gamma = (5.7^{+3.3}_{-1.6}) \times 10^{-10} \text{ second}$$

for the 161-kev transition to the ground state of Sn^{117}. With $\Gamma_0/\Gamma = 0.87$, the mean life of the 161-kev level becomes $\tau_{level} = (5.0^{+2.9}_{-1.4}) \times 10^{-10}$ second.

The partial M1 lifetime is approximately two orders of magnitude longer than the single particle estimate (6). This is, however, in line with other M1 transitions and does not indicate an effect of the l-forbidden character of the transition. The partial E2 mean life is longer than the single particle estimate or at best equal to it.

REFERENCES

(1) R. K. GOLDEN AND S. FRANKEL, *Phys. Rev.*, Vol. 102, p. 1053 (1956).
(2) L. W. FAGG, E. H. GEER AND E. A. WOLICKI, *Phys. Rev.*, Vol. 104, p. 1073 (1956).
(3) See, for instance, F. R. METZGER in "Progress in Nuclear Physics," Vol. 7, edited by O. R. Frisch, New York, Pergamon Press, 1959.
(4) A small correction (~2%) was applied to the actual temperatures in order to take into account the effects of the crystalline binding. See, e.g., W. E. LAMB, *Phys. Rev.*, Vol. 55, p. 190 (1939).
(5) L. A. SLIV AND I. M. BAND, *Acad. of Sciences U.S.S.R. Moscow*, Leningrad (1956 and 1958); issued in the U.S.A. as reports 57 ICC K1, and 58 ICC L1, Physics Department, University of Illinois.
(6) J. M. BLATT AND V. F. WEISSKOPF, "Theoretical Nuclear Physics," New York, John Wiley & Sons, Inc., 1952, Chapt. XII.

V.2
196

THE FRANKLIN INSTITUTE LABORATORIES
NICOL H. SMITH, DIRECTOR

AUTOMATED COORDINATOGRAPH

BY

C. W. HARGENS

Many uses will be found for the new automated coordinatograph developed for Aero Service Corporation, Philadelphia. The development project has transformed the respected Haag-Streit coordinatograph into a completely automatic plotting and measuring instrument. The aim achieved in this instance was to automate the instrument without losing the 0.001-in. accuracy obtainable in its original, manual form.

Layouts on metal or glass with this degree of accuracy and nearly four feet in extent are now made from pre-programmed data. This is the plotting function.

In the reverse sense tabulations of coordinate values may be rapidly produced from aerial photographs, astronomical plates, templates, and any other precise, two-dimensional representation. The dimensional values are at the same time punched on tape so that entry into a computer or other data processing arrangement is simplified.

The equipment (Fig. 1) consists of the coordinatograph table, an electronic console, and the input-output unit. The control equipment was designed so that it could be added to the Swiss instrument without drilling holes or making similar alterations. This permits those who already have instruments to add the numerical control components without extensive modification. In other respects the quality of the electro-mechanical drives was made compatible with the original mechanism. The scheme is remarkably foolproof in operation, as is indicated by the absence of limit switches, etc.

The movement of the X- and Y-axis carriages is produced by means of a novel synthesis of steel bands and short leadscrews. This combination makes it possible to move rapidly over large distances and yet locate precisely to one thousandth of an inch. It is possible also to change driving speeds as much as 10,000 to one. A differential and electric clamp form a part of this mechanism. In the automated instrument speed changing is done automatically, a decade at a time, as the more significant digits reach coincidence with the numerical commands.

The drive system of the new coordinatograph is completely divorced from the distance-measuring or feedback parts. The original rack and pinion which provided the accuracy of the Haag-Streit instrument are retained for their intended purpose. The reading dials have been replaced with highly-resolved digital encoders. These units are capable

FIG. 1. Photograph of equipment.

of reading five decimal digits and therefore can indicate the extremes of carriage motion in either the positive or negative direction.

It will be appreciated that five decimal digits require at least 20 binary digits. There are actually more than this because of lead-lag circuits associated with the most significant digits of the encoders. The conductor system of each moving carriage has nearly 60 circuits including the motor provisions. The digital data are transmitted from the position encoders by means of flat, Mylar sandwich conductors without any sliding contacts or slip rings.

The block diagram (Fig. 2) shows the general arrangement of the feedback and control circuits. The binary-coded, cyclic decimal data from the coordinatograph are translated to Arabic decimal through transistor-buffered relay trees and odd-even complementing circuits. Translations and storage functions are obtained by relay arrays; numerical comparisons and decisions are made by means of diode matrices. There are no vacuum tubes in the equipment and no analog circuits to drift. Pulse counting is avoided on the grounds that a static logic system, directly coupled at all times, is more reliable.

Operation of the comparator circuits is such that decimal digits are

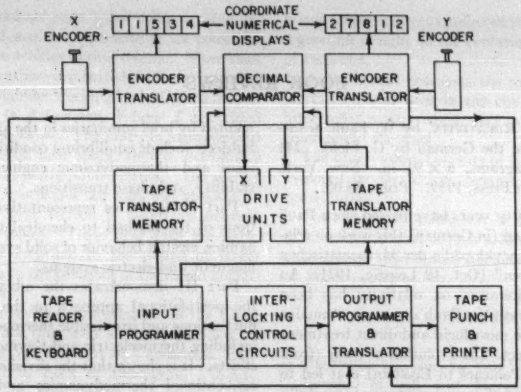

FIG. 2. Block diagram.

examined sequentially. The decimal value of the command is available from the tape translator-storage unit while information in decimal form describing the carriages' location is also sent to the comparator. The most significant digit is compared first, and a "decrease" or "increase" circuit is closed to the high-speed band motor. When coincidence between the command and the carriage location is reached, the next most significant digit's comparator circuit is enabled to function. This process continues, using various drive combinations, until complete coincidence of all X and Y digits is reached. At this instant all motors are stopped, and a prick mark is made.

If the instrument is set for automatic plotting, the memory will be cleared and new data from the tape reader will be entered for the next cycle. The manual controls may be operated at any time without interference with the tape commands. Manual control circuits merely suspend memory control temporarily. When it is desired to again place the memory in control, this may be done without loss of the original entry.

At any time the instrument will print out on demand the carriages' X–Y locations as shown on the numerical displays at the top of the console. If the operator wishes, he may also enter commands directly from the electric typewriter keyboard, without first preparing a tape.

The equipment is designed on a building block principle so that special modifications can be made available for particular applications. A unit for "reading only" is such an example. More elaborate systems include line-following and closed-circuit television scanning equipment. A number of marketable units are now under construction by the sponsor.

THEORY OF RELATIVITY, by W. Pauli, translated from the German by G. Field. 241 pages, diagrams, 6 × 9¼ in. New York, Pergamon Press, 1959. Price, $6.00.

Almost forty years have passed since Pauli first published (in German) this work on relativity in "Encyklopädie der Mathematischen Wissenshaften" (Oct. 19 Leipzig, 1921). An enormous literature on relativity has been published since. Yet Pauli's treatment remains probably the most lucid and direct treatment of this subject. The evolution of ideas (from Lorentz to Poincaré to Einstein) that led to the postulates of special relativity is carefully described. The fascinating historical remarks throughout the book, showing the hesitations and misconceptions of the "founders" contribute to improve the understanding of this subject matter. About forty pages are devoted to general relativity. The last thirty pages describe various theories of elementary particles (Weyl's, Mie's, Einstein's).

The translation of the original German text is followed by a substantial section of notes written by Pauli in 1956, especially for this edition.

HENRI AMAR
The Franklin Institute Laboratories and Temple University

THERMODYNAMICS, by Herbert B. Callen. 376 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $8.75.

This text, written for advanced undergraduate and first-year graduate students, develops a logical theory of thermodynamics by logical deduction from a number of basic postulates. This abstract formulation is shown to result in a great simplification of the mathematics, and helps the reader appreciate the logical consistency and basic simplicity of the subject.

In Part I, comprising about half of the book, the general principles of classical thermodynamics are treated with the help of basic concepts and postulates, and some more advanced mathematical formalism which is supplemented by brief summaries in the appendices. Subjects include equilibrium conditions, processes and thermodynamic engines, system stability, and phase transitions.

Part II illustrates representative applications of the method to chemical thermodynamics, electric behavior of solid systems, and magnetic and electric systems.

Part III demonstrates the advantages of the postulational approach in the study of fluctuations and irreversible thermodynamics, including thermoelectric and thermomagnetic effects. It is shown that the division between conventional thermodynamics and statistical mechanics may be shifted so as to include these subjects within the realm of thermodynamics.

As the author states in his Preface, ". . . thermodynamics is the last of the major classical theories to undergo a postulatory revolution." The advances made possible in mechanics by Lagrange's and Hamilton's formulations, and in electromagnetic theory by Maxwell's equations, justify the hope that the endeavor towards more inclusive abstractions, such as the one presented in Professor Callen's book, will be equally fruitful in the field of thermodynamics. The writing is exceptionally clear, and retains a concrete flavor even in the presentation of the most abstract concepts and interpretation of mathematical relationships. Both as a text and a reference, this book may be highly recommended.

KLAUS CAPPEL
The Franklin Institute Laboratories

VISTAS IN ASTRONOMY, Volume 3, edited by A. Beer. 345 pages, illustrations, 7¼ × 9¾ in. New York, Pergamon Press, 1960. Price, $18.00.

The third in what is called a permanent series of astronomical vistas is now off the press and, while it is a smaller and more specialized volume than its predecessors, it maintains the scientific impact of the initial volume.

The international character of astronomy

in the book is once more demonstrated by having 21 authors representing six countries contribute articles to this volume. More than half the authors are British with the majority of the remainder working in the United States.

Vistas in Astronomy, Volume 3 is divided into two sections. The first—in astronautical jargon—is on near space, that is, astronomy to the limit of the solar system. The second section encompasses the stars and galaxies.

In the first section is found an extreme range of topics. These deal with the rediscovery of Neptune to the design of larger telescopes to the behavior of the recent cometary visitors.

To this reviewer the sections on the Clock Paradox are fully and skillfully handled by Professors W. Cockran of Cambridge, England, and V. Fock of Leningrad, U.S.S.R. Academician Fock points out quite clearly what others neglect to do in the discussion of the Clock Paradox and that is that one body receives an acceleration and, thus, the two frames of reference are not equivalent.

The last four contributions are on the two naked eye comets of recent years. Contained in the Porter contribution is the explanation for the cometary "beard" of the Arend-Roland comet.

Part II of the book is a noteworthy contribution to astronomy. Prof. John B. Irwin presents a challenging picture as to what the astronomer could do were he given basic data on 70,000 stars. He develops the thesis that our basic knowledge of the brighter stars is indeed limited and this must be improved. He indicates that by using the photomultiplier tube in place of the photographic plate, it is possible to reach the 70,000 stars with a new order of accuracy. Prof. Irwin indicates that four observatories distributed in latitude in four to five years could complete this survey. Enough basic data could be accumulated to provide a breakthrough in our knowledge of stellar evolution. Prof. Irwin presents a significant argument and it is hoped that this will be implemented in the near future.

Dr. Olin J. Eggen has contributed one of the most significant articles in the book. Like Dr. Irwin, he is concerned with stellar evolution and he would like to attack the problem from the color-luminosity plots of the stars. Dr. Eggen has undertaken a most challenging problem and this article may provide a springboard for subsequent studies which may throw a good bit of light on the evolutionary history of the stars.

Dr. L. Mestel discusses the role of the galactic magnetic field in the condensation of interstellar matter which may eventually become nuclei for stars.

The observational aspects of cosmology are explored by Dr. D. W. Sciama as the final contribution.

The editors and publishers of this series are to be congratulated in their assembling of the vast talents represented between the covers of *Vistas in Astronomy*.

I. M. LEVITT
The Fels Planetarium

A BEGINNER'S GUIDE TO THE STARS, by R. Newton Mayall and Margaret Mayall. 184 pages, charts, 4½ × 7½ in. New York, G. P. Putnam's Sons, 1960. Price, $2.50.

Today astronomy is news. In tonight's sky you will find a complex of earth satellites and in the newspapers we find hints as to where in the sky to look to see them. With the advent of astronautics there has suddenly appeared a renewed and highly developed interest in the skies. Several books have appeared to aid those with this interest and guide them in learning the constellations.

One such book is *A Beginner's Guide to the Stars*. Written by two highly qualified astronomers, the book is a magnificent primer to use in the introduction to the sky.

The book is divided into two sections. The first deals with what is in the sky such as stars, moon, planets, comets, clusters and, finally, the man-made objects—the satellites. The second half of the book is an array of charts of the constellations with significant information concerning the star group.

Accompanying the map are found the time of visibility of the constellation, its location, the characteristic identifying marks or features of the constellation and, finally, some notes in which we learn whether the constellation is a zodiacal, where are the interesting telescopic objects and other points of interest.

The Mayalls have put together a most readable pocket-sized book which is highly recommended for the beginner.

I. M. LEVITT
The Fels Planetarium

PUBLICATIONS RECEIVED

AEROSPACE DICTIONARY, by Frank Gaynor. 260 pages, 6 × 9 in. New York, Philosophical Library, 1960. Price, $6.00.

DESIGN FOR A BRAIN, by W. Ross Ashby. Second edition, 286 pages, diagrams, 5½ × 8½ in. New York, John Wiley & Sons, Inc., 1960. Price, $6.50.

EVOLUTION: PROCESS AND PRODUCT, by Edward O. Dodson. 352 pages, illustrations, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price: $5.75 (college); $6.90 (trade).

FATS AND OILS, by H. G. Kirschenbauer. Second edition, 240 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $7.00.

SEMICONDUCTOR ABSTRACTS. VOLUME V—1957 issue, edited by C. S. Peet. 449 pages, 8½ × 11 in. New York, John Wiley & Sons, Inc., 1960. Price, $12.00.

FUNDAMENTALS OF SEMICONDUCTORS, by M. G. Scroggie. 160 pages, diagrams, 5½ × 8½ in. New York, Gernsback Library, Inc., 1960. Price, $2.95 (paper).

AIDS TO BIOCHEMISTRY, by S. P. Datta and J. H. Ottaway. Fifth edition, 266 pages, diagrams, 4 × 6¼ in. London, Balliere, Tindall and Cox, 1960. Price, $3.75 (available from The Williams & Wilkins Co., exclusive U. S. agents).

THE STORY OF CHEMISTRY, by Georg Lockemann. 277 pages, 4½ × 7½ in. New York, Philosophical Library, 1960. Price, $4.75.

SCIENCE AND ENGINEERING IN AMERICAN INDUSTRY. Report on a 1956 Survey. 117 pages, 8 × 10½ in. Washington, U. S. Government Printing Office, 1959. Price, $0.70 (paper). NSF 59–50.

STATISTICAL HANDBOOK OF SCIENCE EDUCATION. 94 pages, 6 × 9 in. Washington, U. S. Government Printing Office, 1960. Price, $0.55. NSF 60–13.

THE RESTLESS ATOM, by Alfred Romer. 198 pages, diagrams, 4 × 7 in. New York, Doubleday & Company, 1960. Price, $0.95 ($1.10 in Canada) (paper). Science Study Series S 12.

HORNS, STRINGS & HARMONY, by Arthur H. Benade. 271 pages, diagrams, 4 × 7 in. New York, Doubleday & Company, 1960. Price, $0.95 ($1.10 in Canada) (paper). Science Study Series S 11.

THE SOIL AND ITS FERTILITY, by H. Teuschler and R. Adler. 446 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $12.00.

BOOK NOTES

SELECTED SEMICONDUCTOR CIRCUITS HANDBOOK, edited by Seymour Schwartz. V. p., diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $12.00.

This practical handbook provides the electronic engineer with a reference of over 150 usable, well designed, reliable semiconductor circuits representing the most recent advanced engineering techniques as well as design theory information. A very wide range of subject matter is covered, such as: direct coupled amplifiers, low-frequency amplifiers, high frequency amplifiers; oscillators, switching circuits, logic circuits, a-c to d-c power supplies' converters; small signal nonlinear circuits, and transistorized magnetic circuits. The latest type I.R.E. standards are generally used to provide uniformity throughout the book.

METALLIZING OF PLASTICS, by Harold Narcus. 208 pages, illustrations, 5½ × 7½ in. New York, Reinhold Publishing Corp., 1960. Price, $5.50.

Here is yet another addition to the Plastics Applications Series. Complete details are presented in this handy little book, for carrying out every commercial metallizing process for plastics or other non-conductors. It is the first treatment of the subject to deal with actual production procedures, formulations and techniques for all known metallizing methods. These include the copper film process (developed by the author), the deposition of "electroless" nickel coating, a review of the new molded conductive plastics, "gas" plating, the deposition of thick evaporated films and many others. Numerous illustrations are included showing the latest equipment being used in the newer processes. More than a bibliographic source, this book is essentially a metallizing manual for the plastics, electronics and electroplating industries.

SOME MATHEMATICAL METHODS OF PHYSICS, by Gerald Goertzel and Nunzio Tralli. 300 pages, diagrams, 5½ × 8½ in. New York, McGraw-Hill Book Co., Inc., 1960. Price, $8.50.

This new work, designed primarily for physicists, presents a common set of mathematical techniques useful in such fields as quantum mechanics, acoustics, electromagnetic theory, reactor physics, etc. It is the outgrowth of material developed by the authors for graduate courses at New York University. They have sought to given on mathematicians a well-rounded and thorough indication of the basic concepts involved in the study of linear operators, vector spaces, orthogonal functions and Fourier analysis, Green's functions as tools for the solution of physical problems.

STYLE GUIDE FOR CHEMISTS, by Louis F. Fieser and Mary Fieser. 116 pages, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $2.95.

This book considers the problems facing the writer of a thesis, report, or journal paper and presents helpful suggestions for recognition and avoidance of errors. Positive advice is offered on writing techniques for enlivenment of style without aspiring to a literary level. A scientist can learn to put style into his sentences that will hold the reader's attention. Commentaries on points of grammar and rhetoric, and suggestions on speaking are included as well as information on spelling, abbreviations, hyphenation, punctuation, choice of words and Latin and Greek prefixes.

PROBABILITY: AN INTRODUCTION, by Samuel Goldberg. 322 pages, diagrams, 8½ × 5½ in. New York, Prentice-Hall, Inc., 1960. Price: $7.95 (trade edition); $5.95 (text edition).

This new volume deals fully with the essentials needed by all who use probability concepts, apply statistical techniques, or measure the effects of chance phenomena. Logical in its approach, the book formulates its probabilistic theory by employing the language and notation of sets. By emphasizing experiments with a finite number of possible outcomes, the study requires no previous knowledge of calculus and can be readily understood by the reader with a good background in high school algebra.

GLASS: ITS INDUSTRIAL APPLICATIONS, by Charles J. Phillips. 252 pages, illustrations, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $6.95.

The manufacture, physical and chemical properties, and particularly, the applications of all types of glass, are covered in this volume. It is directed toward engineers, scientists, production and purchasing personnel who are involved in the design, specification, testing, fabrication, purchase or sale of glass products. The author avoids both extreme detail and the popularized "no-data" approach by condensing a mass of information to tabular or graphic form. The book reflects all the recent advances in the field and includes new information on such subjects as glass bead reflectorization and the "floot glass" process for making flat glass. Maximum continuity has been arranged so that each topic is preceded by everything needed to understand it.

PHOTOCHEMISTRY IN THE LIQUID AND SOLID STATES, edited by Lawrence J. Heidt, Robert S. Livingston, Eugene Rabinowitch and Farrington Daniels. 174 pages, diagrams, 8½ × 11 in. New York, John Wiley & Sons, Inc., 1960. Price, $6.00.

Based upon a symposium sponsored by the National Academy of Sciences and the National Research Council, this volume contains the contributions of outstanding authorities in the field of photochemistry. The authors present the basic principles of photochemistry storage, survey the field of photochemical reactions, and state the requirements for reaction types which might prove useful for storing solar energy. Basic research findings are also presented with suggestions of areas for further research which can and will lead to the use of the sun as an important and inexpensive source of energy.

PRINCIPLES OF UNIT OPERATIONS, by A. S. Foust, L. A. Wenzel, C. W. Clump, L. Maus and L. B. Andersen. 578 pages, diagrams, 8½ × 11 in. New York, John Wiley & Sons, Inc., 1960. Price, $15.00.

This book breaks away from the traditional treatment of unit operations concept and offers a new approach that is more economical in time. The unit-operations are presented as unified groups of operations stemming from identical fundamentals. The developments are built up from a simplified physical model or a basic mathematical relation, or both, using generalized notation. After a thorough coverage of the simplified models, the treatment progresses to the more complicated handling of realistic problems by applying the general equations to the specific operations.

ANALYTIC FUNCTIONS, by L. V. Ahlfors, H. Behnke, L. Bers, H. Grauert, M. Heins, J. A. Jenkins, K. Kodaira, R. Nevanlinna and D. C. Spencer. 197 pages, diagrams, 6 × 9 in. Princeton, Princeton University Press, 1960. Price, $5.00.

This book contains one article each by the authors representing a comprehensive survey of recent developments both in the classical and the modern fields of the theory of analytic functions. In the article by Ahlfors, the space of closed Riemann surfaces is shown to possess a complex analytic structure in which the periods of Abelian integrals are analytic functions. A survey of recent advances in the theory of complex spaces is presented by Behnke and Grauert. A proof of the Teichmuller-Ahlfors theorem based on the Beltrami equation is given by Bers. Lindeloef's principle and Martin's positive harmonic functions are the basic elements in a theory developed by Heins. New and old results on the coefficient problems for Schlicht functions are derived by Jenkins by a unified method. Analytic functions on manifolds of two complex dimensions are treated in the article by Kodaira. Nevanlinna presents a simplified proof of Liouville's classical theorem on conformal mappings in Euclidean spaces of dimension 2. Deformations of complex analytic manifolds are studied by Spencer.

THE DYNAMIC BEHAVIOR OF THERMOELECTRIC DEVICES, by Paul E. Gray. 136 pages, diagrams, 6 × 9 in. Cambridge, The Technology Press of M.I.T.; New York, John Wiley & Sons, Inc.; 1960. Price, $3.50.

This volume investigates the small-signal dynamic behavior of thermoelectric devices. Exact analysis of thermoelectric heat pumps and generators is difficult because these devices are described by differential equations that contain product-type nonlinearities. But since the dynamic behavior of these devices is of interest principally as a consequence of the requirement for control of the device, much useful information can be obtained by a small-signal analysis. This approach permits the treatment of linear models that describe behavior of the devices. Analysis of these models produces a number of small-signal transfer-functions that can be used either to compute the response of devices in the frequency-domain, or to calculate the response of devices in the time domain. The results of this analysis can also be used to evaluate the quantitative effect of the materials and device parameters on the dynamic behavior of devices.

V.2
196

INTRODUCTION TO STATISTICAL THERMODYNAMICS, by Terrell L. Hill. 508 pages, diagrams, 6 × 9 in. Reading (Mass), Addison-Wesley Publishing Co., Inc., 1960. Price, $9.75.

This book represents an introduction to equilibrium statistical mechanics, designed as a text for courses offered in chemistry or physics departments on the advanced undergraduate-graduate level. The emphasis is on recent advances in statistical thermodynamics, and consequently the more traditional material is treated in a rather condensed manner to allow ample room for a survey of the important developments of the 1940's and 1950's. In keeping with the author's aim of providing only an *introduction* to the subject, topics are not treated exhaustively and at an advanced level. Thus, although the coverage of subjects is very broad, is it quite appropriate for a beginning text.

CHEMICAL INSTRUMENTATION: A SYSTEMATIC APPROACH TO INSTRUMENTAL ANALYSIS, by H. A. Strobel. 653 pages, diagrams, 6 × 9 in. Reading (Mass), Addison-Wesley Publishing Co., Inc., 1960. Price, $9.75.

Written at the advanced undergraduate-graduate level, this book may be used successfully as a text for a course in instrumental analysis to two semesters in length. The book's most distinctive feature is its attempt to convey a basic appreciation for the design and operation of instruments as a part of the larger scheme of making physical measurements. To facilitate this appreciation, background chapters are presented on the science of instrumentation, error analysis, interaction of electromagnetic radiation and matter, physical optics, electrical electrochemical theory and electronics. Since the author's objective is to develop in the student a genuine understanding of chemical instrumentation, he treats basic techniques fully. A good balance is maintained between physical theory and design on one hand, and chemical theory and procedure on the other. Above all, the text offers a systematic approach to the basic principles common to all methods and types of instruments, rather than a mere catalog of equipment which is rapidly outdated.

CALCULUS OF FUNCTIONS OF ONE ARGUMENT, WITH ANALYTIC GEOMETRY AND DIFFERENTIAL EQUATIONS, by Edward J. Cogan, Robert Z. Norman and Gerald L. Thompson. 587 pages, diagrams, 6 × 9 in. New York, Prentice-Hall, Inc., 1960. Price, $8.50.

This study of calculus, completely in accord with modern views, applies the precise language of mathematics to the exploration of concepts to their deeper roots. Without departing from traditional notation, this book breaks with classical treatment to promote reasoning rather than mere problem solving, and emphasizes the function concept throughout. Employing contemporary terminology, the reader is aided in understanding the difficult and advanced material by the use of such unifying ideas as solution sets, motivation by intuition and application, and integration of calculus with trigonometry and analytical geometry. Because of its independent viewpoint and modern approach, it will be especially useful to readers interested in business administration as well as those specializing in science and engineering.

CHEMISTRY OF NATURAL GUMS AND MUCILAGES AND SOME RELATED POLYSACCHARIDES, by F. Smith and R. Montgomery. 627 pages, diagrams, 6 × 9 in., New York, Reinhold Publishing Corp., 1959. Price, $18.00.

The purpose of this book is two-fold: to evaluate critically the progress made in this fascinating field, and to encourage new ideas for the industrial application of gums. Everything that is known of the chemistry of these versatile substances is presented in this comprehensive new monograph; this includes the rapid strides made in the last few years. The treatment of the subject is exhaustive, and begins with gum function and origin. The gums are classified botanically before the text takes up their chemical and physical properties. Because of the growing commercial importance and possibilities of these raw materials, this monograph provides industry with the sound basis of information it needs for the successful development of gums and mucilages.

NATIONAL BUREAU OF STANDARDS NEWS

IMPROVED METHOD FOR MEASURING ABRASION RESISTANCE OF COATINGS

The National Bureau of Standards has made a number of improvements in its Abrasive Jet Method for measuring the abrasion resistance of protective coatings. These improvements[1] have resulted from continued studies by A. G. Roberts, who originally developed the method in 1955. The changes have not altered the basic manner of operation but have significantly increased the speed, ease, and precision of the measurements.

In the basic method,[2] pressurized gas propels an abrasive powder from a vibrating storage chamber through a nozzle onto the test specimen. Abrasive flow, pressure, distance, and angle are all closely controlled. Abrasion resistance is measured by the time required to abrade through the coating and is expressed in terms of a unit thickness. The first show of bare substrate is the end point and is clearly detected by an abrupt change in color or gloss. The high speed of the particles in the abrasive jet provides a rapid rate of abrasion that permits most coatings to be tested in a matter of seconds. The abrading conditions can be varied over a wide range to permit the testing of virtually any type of coating.

The small scale of operation obtained with the equipment allows the economical use of continually fresh abrasive as well as a large number of replicate tests on a single specimen. Thus the use of a reference panel for calibration and standardization purposes is feasible and has become an important part of the improved procedure. The values recorded for the reference panel also provide a reliable means for reducing data obtained under different conditions to a common basis for comparison. The coefficient of variation for measurements on the reference panel is less than two per cent.

Close control of the amplitude of vibration is necessary to assure a uniform flow of abrasive and a constant rate of abrasion. Recent studies have shown that the amplitude of vibration changes with the gradual drop in the level of abrasive as operation continues, because of a shift

[1] For further details, see "Improved NBS Abrasive Jet Method for Measuring Abrasion Resistance of Coatings," by A. G. Roberts, *ASTM Bul.*, No. 244, p. 48 (1960).

[2] "Abrasive Jet Method for Measuring Resistance of Organic Coatings," by A. G. Roberts, W. A. Crouse, and R. S. Pizer, *ASTM Bul.*, No. 208, p. 36 (1955); and *NBS Tech. News Bul.*, Vol. 39, p. 33 (1955). Apparatus for Measuring Abrasion Resistance, U. S. Patent No. 2,907,200, Oct. 1959.

V.
19

in the center of gravity of the electromechanical vibrator system. Since the change is gradual, it can be compensated by periodic readjustment of the regulated line voltage applied to the vibrator, in order to maintain the standard abrasion rate on the reference panel. Undesirable turbulence and interference effects in the abrasive storage chamber have been minimized by redesigning the chamber so that the abrasive particles sift into and are propelled along a straight tube instead of through a circular chamber. Use of a more free-flowing abrasive (aluminum oxide) than was formerly employed also has improved the uniformity of operation.

New devices for setting and adjusting nozzle-to-coating distance have been added with which distance is quickly and easily set to within 0.002 in. Several subtle sources of error in the original method have been identified and eliminated. Practical operating limits have been established within which good precision can be expected over a wide range of test conditions.

Studies have shown that the method is not limited to the organic coating materials for which it was originally designed. It is also applicable to ceramic, porcelain, and anodic coatings, as well as films and tapes when attached to a rigid substrate. Sheet metal of aluminum, brass, or steel may also be tested; their rates of abrasion are of the same order of magnitude as those of the more resistant types of organic coatings. A good qualitative correlation has been observed between test results and the service performance of protective coatings applied to military aircraft.

The speed and simplicity with which the measurements are made, the reproducibility of results, and the opportunity for calibration and standardization afforded by employing a practical reference panel are features of the improved method that offer a promise of considerable usefulness in the field of abrasion testing.

SPECTRORADIOMETER FOR FLUORESCENT LAMPS

A number of methods are available for evaluating the color of fluorescent lamps, and for studying the effects of the many variables in their manufacture. However, the expanding applications of fluorescent illumination today show that when these lamps are supplied by different manufacturers it is not always possible to achieve a perfect color match. The National Bureau of Standards has therefore constructed a spectroradiometer to calibrate fluorescent lamp color standards that may be used throughout the industry to obtain color uniformity.[3] Measurements made with this instrument will also form a basis for international comparisons of fluorescent lamp colors.

[3] For further technical details, see "Ratio-Recording Spectroradiometer," by H. K. Hammond, W. L. Holford, and M. L. Kuder, *J. Research NBS*, Vol. 64C, p. 151 (1960).

Spectroradiometers generally have three basic components: an element for dispersing radiant flux from a light source into a spectrum, a radiometric detector for measuring this flux in each wavelength, and a recording device for plotting the resulting spectral irradiance data. From these data the color of the source is determined. Some spectroradiometers[4] measure the current from the detector to find the spectral irradiance of fluorescent lamps. In the present work, however, a method was desired that would be essentially independent of both the response characteristics of the detector and the spectral transmittance of the dispersing element. A system was therefore designed by H. K. Hammond, W. L. Holford, and M. L. Kuder of the Bureau staff, in which a test lamp is compared directly with an incandescent lamp of known color temperature.

A double-prism monochromator with a cam-linearized wavelength drive from 360 to 760 mμ is used as the dispersing element in the NBS instrument. The detector is a multiplier phototube with high spectral sensitivity throughout the visible spectrum. A 10-in. strip-chart recorder is used to plot the relative spectral irradiance ratio of the two light sources. These three components, together with the power supplies and electronic circuitry, are housed in a console. The two light sources are mounted on the outside of the console at adjacent entrance ports.

Optical attenuators, in the form of adjustable apertures, are located in the path of each light beam. Each attenuator has two identical components, one fixed and one movable. Each component is a flat metal foil with an array of 25 diamond-shaped holes, 1/10 in. on a side. The two components of an attenuator are fitted together in such a way that they slide over each other to enlarge or decrease the size of the holes as required.

In operation, the spectroradiometer compares the radiant flux from each source as the spectrum is scanned. When the intensities of the two beams are unequal, the unequal phototube signals cause a servo motor to adjust the attenuators, simultaneously opening one and closing the other, until the beam intensities are equal. The attenuation required at each wavelength for equalization is automatically plotted as a function of wavelength over the visible spectrum. The curve thus recorded is the ratio of relative spectral irradiance from the two sources.

Wavelength correlation between monochromator and recorder is maintained by driving the prisms and the strip chart with a second motor. The speed of this motor is varied according to the magnitude of the unbalance signal. When the curve is flat, the rate of scan is maximum; when the curve becomes steep, the rate of scan is auto-

[4] "An Automatic Recording Spectroradiometer," by V. K. Zworykin, *J. Opt. Soc. Am.*, Vol. 29, p. 84 (1939); and "Spectroradiometry," by F. J. Studer and W. R. Jacobsen, *GE Review*, Vol. 52, No. 10 (1949).

matically reduced to assure faithful recording of the peaks of sharp spectral lines. Auxiliary gears and clutches provide wavelength scales of either 1 mμ or 10 mμ for each inch of chart length. At the latter rate, the time required for scanning the visible spectrum is about 8 minutes.

Although the spectroradiometer was built primarily for the calibration of fluorescent lamps, it can also measure the relative spectral irradiance of any source or source-filter combination. Luminescent or fluorescent materials can be measured if they have several square inches of uniform luminance and if the luminance is at least one footlambert.

YOU CAN ADVANCE SCIENCE EDUCATION

Today, more than ever before in its 135-year history, there is vital need for The Franklin Institute to effectively promote education in science and technology. It is imperative that we meet this challenge by providing adequate educational opportunities in these fields. This requires vision, objective planning, and money. We have more than enough of the first two requisites, but far too little of the third.

Our programs are aimed at professional scientists and industry, as well as the lay public and young people seeking inspiration and guidance in choosing a career. The Institute's educational programs are impressive, for they begin with students in the early grades of our elementary schools and continue throughout an individual's professional or industrial life. With more funds at our disposal, the scope and vigor of these activities could be greatly increased.

The Franklin Institute is not richly endowed. It is a non-profit organization, depending for encouragement and support on an understanding public. Capable and conservative management assures wise administration of all funds.

Your gift or bequest, large or small, will be deeply appreciated and will be used effectively to broaden the Institute's educational usefulness. There is a warm satisfaction in giving financial support to an organization that has pioneered in, and is dedicated to, the advancement of science and technology through education.

When property is transferred, title should be in the name of The Franklin Institute of the State of Pennsylvania for the Promotion of the Mechanic Arts.

The Secretary of The Franklin Institute will gladly furnish you with additional information. Write to him at The Franklin Institute, Benjamin Franklin Parkway at Twentieth Street, Philadelphia 3, Pennsylvania.

Language Translation by Machine. —The United States Air Force may have the answer to the age-old problem of international tongue-twisting through its program of language translation by machine. An electronic translator capable of turning Russian into English at the rate of 35 words a second has been developed under a five-year project at the Air Research and Development Command's Rome Air Development Center, Griffiss AFB, N. Y.

The only one of its kind in existence, the translator has been in use on an experimental basis since June 1959. According to Air Force engineers, the English output is "no literary masterpiece" since translation at this point is on a word-for-word basis. However, the machine gives a very good idea of the general content of a Russian article—even though it ignores many of the grammar rules. By the end of this year a sophisticated word analyzer with logic circuits for sentence structure will be added into the translation complex.

Heart of the automatic language translator is a "photoscopic memory" invented by Dr. Gilbert W. King of International Business Machines Research Center, Yorktown Heights, N. Y. A transparent disk only 10 in. in diameter, this dictionary unit can store 550,000 Russian-English words in an area the size of a postcard. They appear around the edge of the disk in concentric tracks of binary code. As a Russian word is fed into the translator via a punched tape, it is "read" by the machine and converted into electrical signals. These are matched in a lightning dictionary drill with the coded equivalents on the glass disk. The English translation is transmitted over an electric typewriter. Any word in the dictionary disk can be located by the machine in less than 1/300 sec.

The Intelligence Laboratory at the Rome Air Development Center is using the research and knowledge of top language scientists in this country and abroad. It has contracts with seven universities and four industrial firms in the automatic language translation program.

The problem of machine translation falls into two essential parts: First, correlation of logic when basic language structures differ as much as they do in Russian and English; and, second, development of the machinery to relay this correlation into electronics. At present there are still some major obstacles in the way of fully-automatic, high quality translation.

For instance, although translation can be completed at extremely high speeds—the translator scans 16,000 words a second—the process of feeding material to the machine is comparatively slow.

The tape containing the Russian text is prepared manually by a typist working with a Russian keyboard typewriter. Using this method, only 40 words a minute are prepared for translation. To overcome this lag, a special input machine is being designed for the Air Force by Baird Atomic Inc., Cambridge, Mass. A Russian print-reader, which will scan the material electrically and automatically make the required tape, will be ready for test later this year. It will bring the input capability up to 2400 words

a minute or sixty times faster than the human operator.

A forerunner of the Russian print-reader which reads English electronically was successfully developed under contract to RADC by Intelligent Machines Research Company in May 1959 and is presently in use.

Another problem—that of transmitting graphic material along with the written word—is now under study for the Air Force by the Syracuse University Department of Graphic Engineering. Design for the construction of a machine to handle this pictorial output will be initiated in May.

The "thinking" portion of the translator—the dictionary and word analyzer—is also being improved at IBM Research. A lexical "buffer" has recently been added which stores words coming from the photoscopic dictionary in complete sentences. The "analyzer" will provide corrected grammar.

While the Air Force's primary interest is in the translation of Russian technical material, the machine's present vocabulary enables it to translate all Russian texts with the possible exception of poetic language where the author may attach different meanings to certain words. If the translator does come across a word not on its glass listing, it transcribes it in red for later inclusion.

The benefits to be gained for both scientific and academic purposes through the use of such a translator are many. At present the U. S. government is spending over a million and a half dollars annually for the job of converting Russian publications into English. Because of the high expenditure of time and money required by human translation, only a part of the necessary work is being done.

It is estimated that the Russians publish a billion words a year—one third of these in scientific fields. The U. S. government translates 80 million words a year at an average cost of two cents a word. Even in its present state of development, the USAF automatic translator— which can be operated by an individual with no knowledge of Russian—is able to render a page of Russian into English in half a minute compared to 45 minutes required by the average human translator.

The Air Force language translation program, started in 1955, will be completed late in 1961 when the experimental model of the fully integrated translation complex is unveiled. Estimates are that by that time five million dollars will have been spent on the project. Air Force engineers and scientists believe this amount is only a drop in the budget bucket compared to the returns that automatic language translation will pay off in the years to come.

Turret System for Helicopters.—A new lightweight turret system for helicopters which will permit accurate cockpit control of an externally mounted machine gun or surveillance TV camera has been developed by General Electric's Missile Production Section, Burlington, Vt., under a U. S. Army Ordance contract. The remote-control system is described as being easily and inexpensively adaptable to existing helicopters and light observation aircraft.

Capable of handling 150-lb. loads, it will accommodate any small machine gun up through 50 caliber to give low-speed aircraft protection against ground fire. It can also be used to provide remote-control for a conventional closed-circuit television camera or General Electric's new Image Orthicon "see-in-the-dark" TV camera for nighttime reconnaissance.

The system can be provided as a kit for installation by second-echelon field maintenance personnel, according to company engineers. Once initial installation is completed, it can be removed entirely or adapted in the field to either armament of surveillance equipment in a matter of minutes to meet specific mission requirements, they said.

Previewed here at the annual meeting of the American Helicopter Society, the system consists of a mount trainable in azimuth and elevation, drive motors, servo amplifiers, and a pistol-type operator hand control. It has no hydraulic components and uses standard 24-volt electric power. Total weight of the transistorized system is less than 80 pounds.

The mount is capable of 370 degrees horizontal rotation and could provide a vertical range of plus or minus 45 degrees. Quickly maneuverable, it provides slew speeds up to 70 degrees per second, the engineers said.

The cockpit control, handled and aimed like a pistol, is suspended from the upper cabin framework by a parallelogram suspension system. This suspension permits the co-pilot to freely move the pistol grip for clear aim from a natural position. The mechanism is stowed out of the way in the cabin when not in use.

In addition to its airborne applications, the system can be employed where unattended control of cameras or similar equipment in remote, inaccessible or dangerous locations is desired. The engineers pointed to possible future uses with closed-circuit TV cameras in plant or prison security systems, at missile launching sites and in radioactive atmospheres.

Rough Terrain Crane.—A "sophisticated" 20-ton capacity rough terrain crane, designed for a variety of applications including operation in four feet of sea water, is now under test at the U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Va.

It is designed for use in military construction projects including clamshell work, dragline work, steel erection, bridge assembly, and other uses which are normal to crane operations. It also will be utilized for barge or landing operations, working out in the surf to assist in unloading landing craft.

Weight of the basic 20-ton rough terrain crane, with a 30-ft. strut boom, is held to approximately 50,000 lb. The crane is expected to be able to travel the highways, under its own power, at speeds of 30 mph., climb 40 per cent grades at a speed of 2.25 mph., negotiate 10 per cent side slopes and, by using maximum front and rear steering, turn inside a circle with a 30-ft. radius.

One of its unique features is a special system which permits the inflation and deflation of the four tires while the vehicle is standing still or in motion. Air pressure can be varied between 10–85 lb., giving added traction when needed in certain types of soil conditions. Also, hydraulic motor-driven outriggers are permanently located on the carrier, and are lowered to help stabilize the crane when lifting maximum loads.

Mounted on a specially designed diesel-driven, off-highway type carrier, the crane is capable of operating without support equipment for sustained periods of time on and off the highway. It is even equipped with a dozer blade, with limited capabilities, to permit the machine to function independently of other construction equipment on the clearing and leveling of its own work site.

The crane on the carrier is mounted on hook rollers and powered by a diesel

engine. Equipped with a 30-ft. strut boom or a 30-ft. gooseneck boom, and three 20-ft. insert sections, it is full revolving. Additional counterweights are provided for operation with a 90-ft. boom. Using the 30-ft. boom and outriggers, the unit can lift 10 tons at a 20-ft. operating radius. It has a 25,000-lb. drawbar pull.

The prototype crane, currently undergoing tests, was built by LeTourneau-Westinghouse Company, Peoria, Ill., under contract with the Laboratories. The superstructure was manufactured by American Hoist and Derrick Company.

Improved Solid Lubricant.—A solid lubricant with as much as 20 times the normal load-bearing capacity has been announced by Dr. Guy Suits, General Electric Vice President and Director of Research. Solid lubricants make possible the miniaturization of many devices, since they can be used at higher temperatures than the liquid type, Suits noted. They also provide permanent lubrication, improve reliability, and reduce maintenance.

The improved lubricant was achieved by the addition of various inorganic sulfides to the well-known lubricant molybdenum disulfide, or to tungsten disulfide, Suits revealed. The effect of the additives was discovered by Dr. Arthur J. Haltner and Curtis S. Oliver of the General Electric Research Laboratory, and was evaluated in a long series of tests carried out in collaboration with other staff members.

When a ½-in. diameter chrome alloy steel hemisphere slid over a mild steel surface, using molybdenum disulfide as the lubricant, the lubricant film was able to withstand a one-quarter kilogram load. When an additive (silver sulfide) was incorporated in the lubricant, the load-bearing capacity increased to 5.3 kilograms. Even better results are obtained when the lubricants are used on relatively hard surfaces, such as chromium.

The additives are all inorganic sulfides, and their effects can be noted at relatively low concentrations. During the tests conducted thus far, they were typically added in quantities of approximately 10 per cent by weight. Among the most effective of those that have been investigated are the sulfides of antimony, platinum, mercury, silver, titanium, and lead. When used to lubricate iron or steel surfaces, such additives apparently decompose to produce a layer of iron sulfide on the metal surface. This layer is believed to have better lubricating properties than the original metal, and may interact advantageously with the lubricant itself. The mechanism by which the additives improve lubrication may be similar to that which is effective when "extreme-pressure" additives are used in liquid lubricants.

Studies carried out by General Electric indicate that chemical reactions between molybdenum disulfide and the surrounding atmosphere may have a significant effect on lubricating efficiency, to an extent not previously appreciated. The investigations of molybdenum disulfide are part of an extensive program being conducted at the GE Research Laboratory on the fundamental mechanisms involved in lubrication by solids.

Inexpensive Method for Reproducing X-Ray Films.—A recent issue of *The University of Michigan Medical Bulletin* reports that doctors and a photographic specialist at the U-M have found a simple inexpensive method for reproducing X-ray films with "sufficient detail to maintain diagnostic quality." Previous methods have been expensive and time-consuming, and have rarely produced satisfactory

copies. *The Bulletin* says the new development is highly important in getting maximum distribution of unusual X-ray pictures for medical consultations and teaching purposes.

The account was written by Dr. Walter M. Whitehouse (M.D.), associate professor of radiology; Fred Andregg, supervisor of U-M Photographic Services; and Dr. Fred J. Hodges (M.D.), chairman of the U-M Department of Radiology.

They employed an improved diazo method of reproduction with a standard diazo machine—somewhat like a blue-printer—used for reproducing documents.

However, the resulting duplicates differ from blue-prints in that they have a transparent base and a "direct positive continuous tone." A new type of diazo film produced by the Ozalid company made possible the application to X-rays.

The process produces a dry finished copy in two minutes. It differs from the original only in having a thinner film base and a blue tint.

Computer-Directed Navigation System.—A computer-directed map projection system under development by IBM engineers may enable the pilot of a supersonic plane to tell his location over the earth by merely glancing at a screen on his instrument panel.

The experimental system can project a highly accurate circular map of an area 400 miles in diameter (125,000 square miles) on a 7½-in. screen. An airborne computer performs all the calculations necessary to position the map on the screen and pinpoint the pilot's location. He can literally see on the display map what city or town he is flying over.

Heart of the navigation display is a glass hemisphere about 6 in. in diameter. A detailed map of half the Earth is reproduced photographically on its inside surface. A beam of light illuminates a small section of the map, which is then projected onto a flat, translucent screen in front of the pilot. As the plane moves, the computer automatically adjusts the map presentation.

Engineers at the IBM Federal Systems Division laboratory at Owego, N. Y., say that a similar display device could be used in space navigation.

To use the system in space, a technique known as stellar map matching might be employed. A stellar map, or "chart of the heavens," would replace the map of the Earth inside the glass hemisphere and be projected on the screen. Then a telescope view from the space craft would also be projected on the screen, superimposed on the space map projection. When the stellar map and the telescope view are lined up, the computer can automatically determine the position of the space craft.

IBM engineers are now working towards refinement of the system.

Epitaxial Diffused Transistors.—Major reductions in switching times and collector resistances of diffused base transistors have been realized at Bell Telephone Laboratories by the use of epitaxially grown films. These improvements have been achieved by combining diffused base technology with the epitaxial film technique. Films themselves have also been under investigation in a number of other laboratories.

Application of this technique has resulted in a reduction of the switching time of silicon devices by a factor of more than 10, and in a comparable reduction in the collector resistance. This major development is expected to have far-reaching implications in

both the fabrication and application of semiconductor devices.

Diffused base transistors require a relatively high resistivity collector region in order to attain low capacitance and high voltage breakdown. Typically, to afford ease of mechanical handling, this region has been much thicker than required electrically. The excess thickness increases the collector resistance and, through carrier storage, the switching time.

Ideally the thickness of the lightly doped collector region should be in the neighborhood of 0.1 mil, which is about a factor of 30 thinner than normally used. Semiconductor wafers prepared by conventional methods become extremely difficult, if not impossible, to handle as they approach this desired thinness. The problem can be solved by the use of epitaxial film techniques. Lightly doped epitaxial films are grown on and supported by a low resistivity substrate giving the desired combination of electrical properties and mechanical strength. The term "epitaxial" means that the film is a direct extension of the single crystal structure of the substrate wafer.

Diffused base transistors have been made at Bell Telephone Laboratories on epitaxial layers of both germanium and silicon and have exhibited the theoretically predicted improvements. For example, in two similar silicon transistor structures, one conventional and the other using the epitaxial material, switching time in a typical circuit has been reduced from 200 to 20 millimicroseconds. Further, collector series resistance of the epitaxial transistors was reduced by a factor of more than 10 and was comparable to that of conventional devices 15 times larger.

Experiments with germanium indicate that use of epitaxial layers will extend the frequency response of germanium transistors well beyond that of the two kilomegacycle device recently described by Bell Telephone Laboratories.

The new technique involves modification of the starting material and can be introduced directly into the front end of existing diffused base production facilities without any further changes. Single crystal wafers of heavily doped material are first cut and polished. A lightly doped thin film (about 0.1 mil) of the same conductivity type is then deposited epitaxially on the wafer surface. This film provides the desired thin, lightly doped collector region. From this point on, standard diffused base techniques are employed.

The use of epitaxial material in diffused base transistors not only results in major improvements in switching time and collector resistance, but in addition simplifies the design and understanding of transistor devices and brings them closer to ideal forms, such as n-p-i-n structures. Further, the addition of the epitaxial film technique to the well established diffusion technology provides the device engineer with an extra degree of design freedom which should result in new devices difficult or impossible to achieve by older techniques.

Water Supply in Space.—Research for a solution of the water supply for man on an extended space trip is underway at General Electric's Space Sciences Laboratory. The work is being done under a $94,000 contract recently awarded by the National Aeronautics and Space Administration to the Company's Missile and Space Vehicle Department.

Using a unique vacuum pyrolytic catalysis technique, the approach involves vaporizing the volatiles in the biological wastes at reduced pressures, hence low temperatures. The vapor

is then pyrolyzed over a catalyst and thence condensed back to a liquid state. This condensate has been analyzed and found to be without odor, color or toxic materials and adjudged suitable for human consumption, according to General Electric scientists. The technique does not require chemical additives or filtering.

Rats and a monkey will be used in the study to determine of the feeding of recycled recovered water has any effect on animals. The rats will be studied through more than one generation, while the monkey will be carefully observed during a prolonged confinement in a closed area.

Besides visual observation, periodic tests of waste samples will be made to determine the animal's health. If indicated, the monkey will be removed for more thorough physical examination.

The monkey will be provided with a three-day supply of water in the semi-closed system. Dry food and dry air will be added as required. (In a closed system nothing is added or subtracted.) Waste products, including moisture from the vented air will be collected, processed, and returned to the system.

The new approach on water recovery is being investigated owing to the weight and storage problems involved in supplying water by conventional methods, according to General Electric scientists.

The water supply problem is only one of many aspects of man's environment in space now being studied at the General Electric Space Sciences Laboratory. These investigations include research on waste management, oxygen recovery, air purification and food requirements.

V.2
196